ACLGuard: Physics-Aware Knee Loading Monitoring System for Anterior Cruciate Ligament Injury Prevention Training

BAICHEN YANG, The Hong Kong University of Science and Technology, Hong Kong SAR
XINYI ZHANG, The Hong Kong University of Science and Technology, Hong Kong SAR
XIN HE, The Chinese University of Hong Kong, Hong Kong SAR
CHI XU, The Hong Kong University of Science and Technology, Hong Kong SAR
WENTAO XIE, The Hong Kong University of Science and Technology, Hong Kong SAR
ZURU LIANG, The Chinese University of Hong Kong, Hong Kong SAR
PATRICK SHU-HANG YUNG, The Chinese University of Hong Kong, Hong Kong SAR
QIAN ZHANG*, The Hong Kong University of Science and Technology, Hong Kong SAR

Anterior cruciate ligament (ACL) injuries are common in sports and significantly affect athletes' health and performance. Integrating knee adduction moment (KAM) biofeedback into ACL injury prevention training has been shown to effectively reduce injury risk and enable athletes to safely engage in high-risk activities. However, current motion capture-based monitoring methods are impractical for on-field use due to their bulky setups and limited coverage. While Inertial Motion Unit (IMU)-based methods address some of these issues, their poor performance during high-risk tasks limits their applicability in real-world scenarios. This paper presents ACLGuard, a novel physics-aware KAM monitoring system designed for out-of-lab ACL injury prevention training. ACLGuard utilizes a combination of continuous monitoring with a set of IMUs and a one-time body capture with RGB-D camera. We identify key limitations in existing approaches, including *insufficient body information and inadequate encoding of biomechanical principles*. To overcome these challenges, we introduce a one-time RGB-D registration scheme to capture comprehensive body information and develop an inverse dynamics (ID)-guided modeling algorithm to incorporate biomechanical principles into the system. However, extracting kinematic features under high-risk conditions and obtaining representative body features with respect to ID principle are challenging. Even worse, these extracted imperfect features increase the ID-guided modeling difficulty for KAM estimation. To derive meaningful physical features, we propose a hybrid deep learning model referring to motion patterns and physical priors. For ID-guided modeling, we introduce an attention-enhanced multi-task learning framework to establish hidden physical mappings from imperfect features to KAM. We collect a dataset from 10 athletes and 9 non-athlete subjects, containing four main high-risk tasks in real-world ACL injury prevention. Evaluations show that ACLGuard achieves an average root mean square error of 0.176 Nm/kg and a normalized root mean square error of 11.5% in KAM estimation, comparable to existing markerless motion capture solutions but offers an on-field monitoring potential with a significantly simpler setup.

*Qian Zhang is the corresponding author.

Authors' Contact Information: Baichen Yang, The Hong Kong University of Science and Technology, Hong Kong SAR, byangak@cse.ust.hk; Xinyi Zhang, The Hong Kong University of Science and Technology, Hong Kong SAR, xzhangfz@cse.ust.hk; Xin He, The Chinese University of Hong Kong, Hong Kong SAR, xinhe@cuhk.edu.hk; Chi Xu, The Hong Kong University of Science and Technology, Hong Kong SAR, cxubs@cse.ust.hk; Wentao Xie, The Hong Kong University of Science and Technology, Hong Kong SAR, wentaox@ust.hk; Zuru Liang, The Chinese University of Hong Kong, Hong Kong SAR, zuruliang@cuhk.edu.hk; Patrick Shu-hang Yung, The Chinese University of Hong Kong, Hong Kong SAR, patrickyung@cuhk.edu.hk; Qian Zhang, The Hong Kong University of Science and Technology, Hong Kong SAR, qianzh@cse.ust.hk.

ACM 2474-9567/2025/12-ART231
https://doi.org/10.1145/3770692

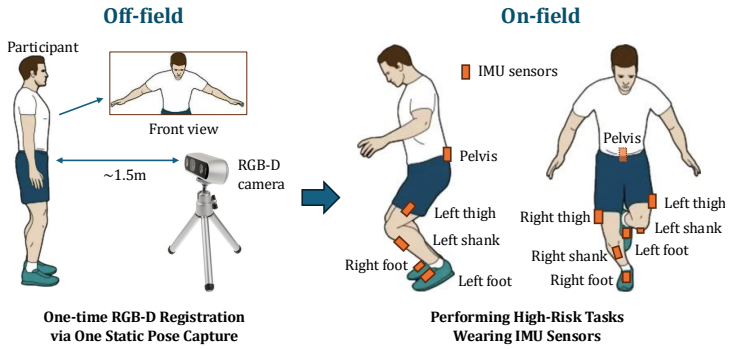


Fig. 1. Usage Scenario.

1 INTRODUCTION

ACL injury is one of the most common and severe knee injuries across sports with more than two million cases worldwide annually [49]. Given the high surgical cost and negative influence on athletes' career [24], prevention training programs have been proposed to mitigate its risk [19, 20, 51]. Since ACL injuries often occur in **high-risk tasks**, which are highly dynamic movements like cutting and pivoting, training programs try to educate athletes to perform tasks in a safe manner by improving neuromuscular control [4, 19, 24].

Evidence has shown that augmenting prevention programs with bio-feedback on risk factors is essential to the athlete's neuromuscular control ability and thus greatly reduces injury risks [19, 24, 29]. The **knee adduction moment (KAM)**, which measures the knee joint loading, is considered a primary ACL injury risk factor and an effective bio-feedback metric [7, 19, 55]. Reducing KAM during training can achieve more effective injury prevention.

Accurate KAM measurement requires a biomechanical analysis system combining motion capture (MoCap) and force plates [15], but the complex setup and needs of on-body reflective markers limit its real-world deployment. Markerless MoCap systems simplify the setup requirements [21, 58]. Though achieving accurate estimation, these systems cannot apply for on-field training due to limitations in field-of-view and occlusion. Mobile sensing solutions are better suited for on-field scenarios, including pressure insole-based [30] and IMU-based [11, 54] ones. While comfortable, current pressure insole-based solutions lack durability under constant heavy loads in high-risk conditions. IMU-based methods are more stable and achieve effective estimation in daily conditions [27, 59]. However, recent IMU-based work performs badly in high-risk tasks [54].

We dive into details to find out the main obstacle hindering the performance of IMU-based methods among high-risk tasks. The key limitation is **insufficient body information and inadequate encoding of biomechanical principles** in modeling KAM. As a physical quantity represents torque, KAM's computation requires not only motion information, but also body information including mass and moment of inertia. Moreover, as is discussed in Sec. 3, the mapping from motion to KAM is highly non-linear and complex, which involves sophisticated physical constraints. Current IMU-based solutions only utilize the motion information and treat the mapping as a black box in an end-to-end manner without much physical prior. This critical insufficiency results in current inferior performance in KAM estimation. From the literature, we notice that we can **complement the body information via one-time RGB-D camera scanning** [30], which can be achieved in an off-field manner. In addition, **leveraging the principle of inverse dynamics (ID)** can better decipher the complex mapping to KAM and guide the modeling [22].

Fig. 2. System Overview.

In light of this, we present ACLGuard, a physics-aware KAM monitoring system for ACL injury prevention. Fig. 1 illustrates our system setup. To provide body information without on-field camera capture, we introduce one-time off-field RGB-D registration. We utilize seven IMUs attached to lower limbs for on-field motion monitoring. We additionally design modeling algorithms based on the ID principle. Fig. 2 shows our system overview. Two main challenges need to be addressed to realize such a system:

- **Extracting kinematics and body features.** Estimating kinematics features from IMUs is difficult in high-risk tasks due to fast and abrupt body changes. Moreover, to cover the high variations of movement patterns, a large amount of data is required. In addition, from the ID principle, the body features we needed are essentially the mass matrix, which is motion-dependent and requires continuous monitoring by motion capture. Extracting representative features from one-time off-field RGB-D capture to provide information regarding mass matrix is challenging.
- **Effective ID-guided modeling on imperfect features.** Given the real-world measurement constraints, the above features might be imperfect and inaccurate, making it difficult to use physical equations explicitly. Moreover, large KAM variations across tasks and subjects caused by motion and landing differences impose difficulty in data-driven methods. Hence, designing a modeling scheme with physics-awareness under such limitations needs additional considerations.

Towards the first challenge, we propose a Biomechanical Information Retrieval module with hybrid deep learning architecture. Specifically, in response to the fast and abrupt motion changes, we design a hybrid neural network architecture with two parallel branches, capturing both short-term motion features and general kinematics patterns. To overcome the data scarcity issue in high-risk conditions, we propose to pretrain the module on large open-source motion capture datasets with synthesized IMUs. In terms of the body features, we identify the representative quantities, namely Body Segment Parameters (BSPs) including segment mass, center of mass and moment of inertia. We then utilize BSPs to supervise the extraction of the most representative body feature toward mass matrix with respect to ID principle analysis from the one-time RGB-D registration.

For the second challenge, we design a Physics-aware KAM Modeling module, which encodes the ID principle into the models. Inspired by the three necessary physical quantities in the ID principle, we design a multi-task learning framework with auxiliary tasks on Mass Matrix, Center of Pressure and Ground Reaction Force to prevent potential error accumulation of explicit physical equation usages on imperfect features. We enhance the framework with a multi-head self-attention mechanism to better fuse the features learned from these tasks. Toward the highly variable KAM patterns in ACL injury prevention training, we additionally enhance the framework with a knee adduction angle-guided Mixture-of-Expert (MoE) design to better capture these variations.

We collaborate with an ACL injury research center and enroll 10 athletes as well as 9 non-athlete subjects, comparable to existing dataset size in related ACL injury prevention works [18, 54]. We collect data with four typical types of high-risk tasks (cutting, pivoting, double- and single-leg jump-landing) in ACL injury prevention training [4, 24, 55]. We closely work with biomechanical experts in high-risk task selection, the correctness and safety of task execution. Our dataset includes variations in knee movements when performing tasks, making it closer to real prevention training scenarios. Evaluations demonstrate the effectiveness of ACLGuard, with an average root mean square error (RMSE) of 0.176 Nm/kg and a normalized root mean square error (NRMSE) of 11.5%, largely outperforming previous IMU-based solutions by reducing over 60% error and comparable to

the Markerless MoCap solution but with on-field monitoring potential and a much simpler setup. The accuracy we achieved is regarded as a reasonably good level in terms of joint moment estimation in related work in the biomechanical field [31].

We summarize our contributions in three folds:

- We introduce ACLGuard, a novel physics-aware KAM monitoring system for on-field ACL injury prevention training. Our system enhances IMU sensing with one-time RGB-D registration method, enabling physics-aware kinetic analysis with motion and one-off visual feature inputs.
- We propose a novel physics-aware KAM modeling framework inspired by the principle of inverse dynamics, achieving effective physical principle encoding into data-driven models based on attention-enhanced multi-task learning. We additionally designed biomechanical information retrieval module to improve the kinematics and inertial feature extraction.
- We evaluate ACLGuard using our collected dataset consisting of 10 athletes and 9 non-athlete subjects. The evaluation results demonstrate an average RMSE of 0.176 Nm/kg and NRMSE of 11.5%, achieving comparable results to markerless motion capture solutions but with on-field monitoring potential and a much simpler setup. The code is open-sourced in the following link: https://github.com/aclguard/ACLGuard

2 BACKGROUND

2.1 ACL Injury Prevention

The ACL is a crucial component of the knee joint, playing an essential role in resisting anterior tibial translation and internal-external rotation of the tibial to ensure knee stability. Under abrupt and high-intensity pulls, the ACL may get injured or even torn up. ACL injuries occur most often during sports that involve sudden stops, directional changes, jumping or landing, such as soccer, basketball, and handball, where the ligament's ability to resist extreme loading is challenged [5].

To prevent ACL injuries, various programs have been proposed to train athletes to reduce risk factors. Among them, augmented training with biofeedback has been proven effective in reducing risk factors. This kind of program will monitor essential biomechanical metrics during athletes' movement, and feedback to them for posture correction to prevent potential injuries. Multiple feedback metrics have been validated for ACL injury prevention [5, 19, 23, 41], among which KAM is considered as a better choice since it has been strongly recognized as the main risk factor of ACL injuries [40, 53, 60].

In terms of feedback scheme, studies [6, 41] have shown that with real-time feedback, the biomechanical metrics of athletes are improved when they perform squatting. Compared to kinematics feedback, kinetics feedback can significantly better help the athletes adjust KAM to the target safe region [19]. Moreover, to reduce KAM, individualized methods are suggested to be tailored for different athletes, as the response of KAM to instructions could diverse a lot among different subjects. With customized feedback and corresponding instruction, instant reductions in KAM are observed in [5], ranging from 13.4% to 17.1%.

Given the importance of ACL injury prevention and the effectiveness of KAM biofeedback training, we are motivated to design a system that can monitor KAM during training activities for ACL injury prevention.

2.2 Knee Adduction Moment

Knee adduction moment (KAM) is the moment of the knee joint in the frontal plane, which is widely used to indicate the loading of the knee joint. It is considered a primary risk factor for ACL injury, as studies have shown a significant correlation between injury risk and KAM value during high-risk tasks like jump-landing and cutting [26, 42, 45, 55]. The potential reason behind is that ACL is not well-equipped to resist adduction moments, as its primary function is to prevent anterior tibial translation and rotation. If KAM exceeds the capacity of the stabilizers, ACL will have a high risk of rupture [60].

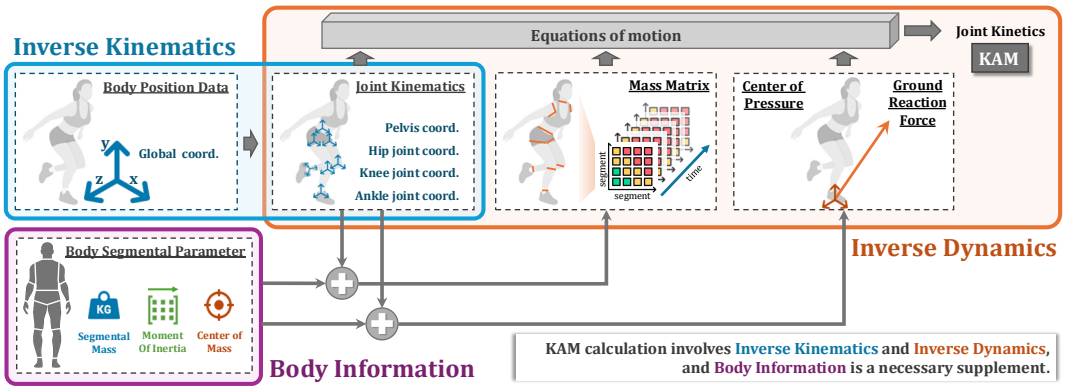


Fig. 3. Design Rationale Illustration. ACLGuard combines body information and joint kinematics to estimate KAM based on learning key components in inverse dynamics principle.

The measurement of KAM is a non-trivial task, as it not only correlates with joint motions, but also biomechanical structures of individual subjects. The fundamental principle to calculate KAM is inverse dynamics, which requires body anatomical measurement, high-precision motion tracking and external force measurement [15]. The detailed inverse dynamics principle will be illustrated in the Sec. 3. The typical way of obtaining these information is through 3D motion capture and ground reaction force measurement system. However, its high setup complexity and measurement area limitation hinders the usage in real on-field monitoring scenarios. We therefore try to provide a more convenient measurement of KAM while preserving the accuracy in out-of-lab training circumstances.

3 DESIGN RATIONALE

To compute joint moments like KAM, the current biomechanical analysis goes through the procedure of inverse kinematics and inverse dynamics [15], which is shown in Fig. 3. Our system's design is largely inspired by these two principles.

3.1 Inverse Kinematics

In detail, such analysis procedure starts with inverse kinematics, which resolves human locomotion in joint space through decoding movements in the Cartesian space based on the kinematic chain of musculoskeletal models. The main idea of inverse kinematics is to solve the corresponding joint angle based on the coordinates of the musculoskeletal system's end effectors in world space. Each joint angle can be determined based on its successor's position and overall understanding of the kinematic chain. The procedure of inverse kinematics is a complicated non-linear mapping [15].

Given the fact that inverse kinematics procedure is highly non-linear and requires prior knowledge of the kinematic chain, in our system design, constructing the mapping from motion to joint space kinematics is a non-trivial task. Inspired by the chain architecture and related work [64], we utilize a deep learning-based feature extractor with hierarchical GRUs to predict joint angles based on their position in the kinematic chain. And as the end effectors' positions, which are the ankle joints in our case, are of paramount importance in constructing the whole kinematic chain, we propose to predict contact information additionally to improve the overall kinematics feature extraction. Moreover, in the ACL injury prevention scenario, the short-term changing patterns of joint

kinematics are essential in capturing peaks of KAM, we propose a CNN-based feature extraction as well to better capture changes in a local perception view.

3.2 Inverse Dynamics

After obtaining kinematics in the joint space, joint moments should then be computed based on the joint space motion as well as external force measurements with respect to the following Newtonian dynamic equation [34]:

$$M(q)\ddot{q} + C(q,\dot{q}) + G(q) = \tau + J_F^T F \tag{1}$$

where $q, \dot{q}$ and $\ddot{q}$ are the joint space kinematics, and M, C and G are correspondingly mass matrix, Coriolis/centrifugal and gravitational effects. The Coriolis and gravitational terms are matrices that describing the earth rotation and mass's impact on the body segments. These terms are relatively stable in daily conditions. On the right-hand side, τ stands for the joint moments and $J_F^T F$ term presents the impact of external force on the system, in our case, the Ground Reaction Force (GRF). Note that this is a segmental equation so multiple equations will be established in the human musculoskeletal system. For a human body system with N degree of freedoms (DoFs), we will have N equations here. In addition, we can have 3 more equations to model the human body as a whole by:

$$M\mathbf{a} = \sum_i \mathbf{F}_i^{ext} \tag{2}$$

where $\mathbf{a}$ is the acceleration of the system root in Cartesian space, and the right-hand side is the sum of all GRFs. From the above equations, we can see that we need to model not only the motion in the joint space, but also the body inertial properties of the musculoskeletal system (mass matrix) for KAM modeling. While the GRF term also makes an influence, there are some redundancies as we have N DoFs but $N+3$ equations in total, where the three more equations come from Eq. 2 that describing external force's impact. If we only get one leg contacting the ground, which is common in highly dynamic conditions, we can obtain joint moments only with information on joint motions and body inertial properties. And for double-leg contacting conditions, some works are reporting a strong correlation between motion pattern and GRFs [48], making the inference of force information from motion information possible.

Inspired by the above equations, we propose to obtain body inertial properties and leverage the inverse dynamics principle to bridge the information and prior gap in existing IMU-based methods. However, the body inertial properties needed in the equations of motion is the mass matrix, which is a motion state-dependent variable. Without continuously monitoring the body movement as well as skeletal structure, it's hard to accurately and directly obtain this feature. By exploring the definition of mass matrix [33], we find that the concept of mass matrix can be disentangled as joint-space motion-related features as well as Body Segment Parameters (BSPs). The BSPs are motion state-independent variables including segment mass, Center of Mass (CoM) and moment of inertia, which can potentially be captured through only one-time body measurement. We therefore propose to capture BSPs for body inertial information, fulfilling the requirements of ID without a complicated setup.

In the meantime, since we are not able to obtain absolutely precise kinematics and body inertial information, and also need to infer external force from motion, it will be quite inaccurate if we try to solve the physical equations directly due to error accumulation. Motivated by the above-mentioned ID principles and the recent success of deep learning-based physics modeling [10], we design physics-aware multi-task learning architecture for final KAM regression. According to the illustrations above, we can see that both the mass matrix term and the GRF term are essential components in the computation of KAM, but we do not have direct access to these features. Notice that both of these components have a strong correlation with joint motion as well as body inertial properties in principle, which may be modeled through deep learning, we therefore design auxiliary tasks with respect to these two terms for better encoding of ID principle into our framework.

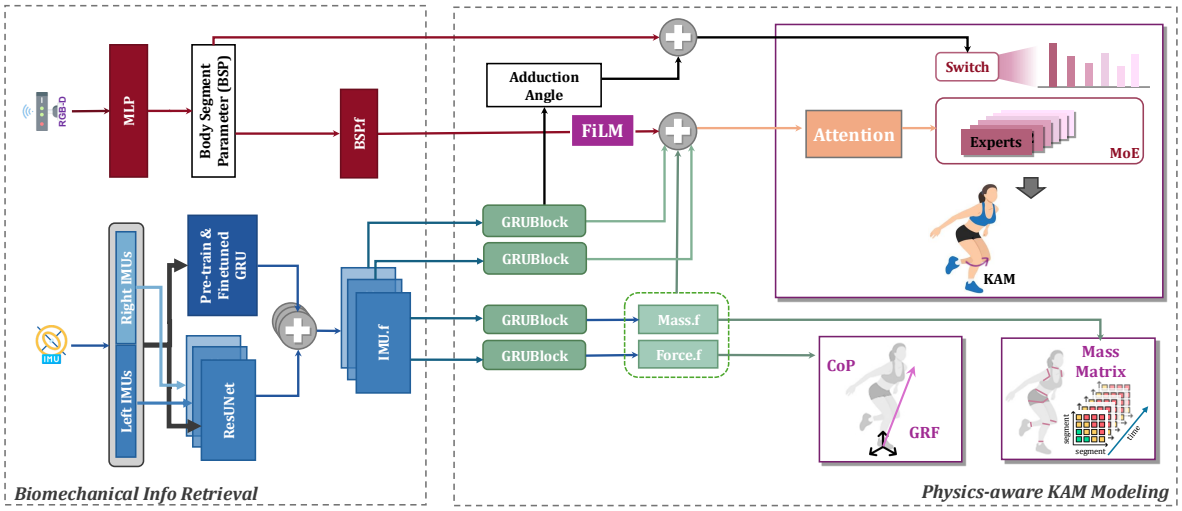


Fig. 4. System Workflow. The left part is Biomechanical Information Retrieval module and the right part is the Physics-aware KAM Modeling module. IMU.f represents the kinematic features extracted from IMUs. Mass.f and Force.f stand for intermediate features that utilized for Mass Matrix regression and GRF, CoP regression.

4 SYSTEM DESIGN

Next, we will introduce the detailed designs of ACLGuard. It mainly consists of two parts: **Biomechanical Information Retrieval (BIR)** module and **Physics-aware KAM Modeling (PKM)** module. The workflow of the system is shown in Fig. 4.

4.1 Biomechanical Information Retrieval

To achieve good KAM monitoring, we need to provide comprehensive and accurate biomechanical features for completing the dynamics analysis. As we have discussed in Sec. 3, we should capture both joint kinematics as well as BSP for inverse dynamics. In a real-world ACL injury prevention training scenario, acquiring joint kinematics and BSP accurately faces difficulties, leading to our following designs.

4.1.1 Kinematic Features Extraction. Getting accurate joint kinematics from inertial sensors is challenging given the difficulty in understanding the human lower-limb kinematic chain from a Cartesian space perspective. Especially, in a highly-dynamic movement scenario, the mapping from sensor readings to kinematics will be extremely nonlinear and complex.

The reasons for kinematics feature extraction difficulty in highly dynamic conditions are in two folds: Firstly, the general joint movement patterns are difficult to be captured correctly due to the frequent contact pattern changes in highly dynamic conditions. Secondly, joint kinematics in ACL injury prevention conditions include many fast and abrupt changes, requiring estimations on not only general body movement pattern, but also short-term and abrupt body changes.

In light of this, we propose a hybrid RNN-CNN model to capture both the general joint angles pattern information as well as subtle and short-term joint kinematic changes as our feature for KAM prediction. Since our main target is knee adduction moment, which is majorly determined by lower limbs, in our design we focus on capturing lower body kinematics. Inspired by [64], we adopted a hierarchical GRU architecture to capture

the general lower limb joint kinematic patterns. Specifically, we design the module with consecutive 4 blocks of 2-layer GRUs, with each block predicting the segmental kinematic information along the kinematic chain. The architecture is shown in Fig. 5. The rationale is that, to calculate upper level kinematic along the kinematic chain, we need to first solve the lower level kinematic, then combine it with motion signal captured on the upper-level node based on the inverse kinematic principle [15]. The first GRU block will be used to estimate leaf node kinematics, which are two ankle joints. Then the second GRU block will regress both leaf and middle joints along the kinematic chain, namely ankle, knee and hip joints' kinematics, and the third GRU block will integrate all joints' information and try to predict root (pelvis) motion. In addition to general motion estimation, we use the last GRU block to predict whether the subject is contacting the ground or not for better end effector position analysis. Since the contacting condition is correlated with body inertial features like segment center of mass, we here add inertial features in addition, which will be illustrated later, for better prediction of the contact information.

Besides general pattern estimation through hierarchical GRU, we additionally design a ResUNet1D-based feature extraction branch to capture short-term motion variations [17, 50]. Though GRUs are powerful in terms of time series feature capturing, they may be insufficient in modeling some of the features within a local perception field [13]. To provide a better perception of short-term changes on the motion time series, like sudden directional change or joint instability, which are common in high dynamic training scenarios, we design a 4-layer UNet1D-like backbone for such feature extraction. The reason for using UNet instead of traditional CNNs is that the jumping connections of UNet can preserve more details of raw inputs while maintaining powerful feature extraction ability given by deep network structure, making the model more sensitive to subtle motion changes [35]. To overcome overfitting issues and empower better convergence, we additionally add residual connections into our UNet architecture.

Unlike general movement patterns, the impact of short-term motion variations are mostly joint-specific. For instance, the major influence of subtle left lower limb variation will be on left knee adduction moment, while general posture of both legs determines the overall pattern of KAM. Therefore, we split three ResUNet blocks with left, right leg and both legs IMUs as inputs to extract features with different emphases on short-term motion variations of different body parts. For left/right knee adduction angle and moment estimation, we fuse GRU feature with features extracted from ResUNet branch with the corresponding leg's IMUs to let the model focus on detailed motion of the that leg. As for GRF, CoP and Mass Matrix regression, the ResUNet feature we used is from the branch with both left and right leg's IMUs, in order to provide a more balanced focus on both lower limbs. For better features capturing from signals, we use the Extended Kalman Filter on accelerometer data and gyroscope data to generate rotation matrix as additional inputs into both GRU and ResUNet branches.

To empower our model with a better general joint kinematics estimation ability, we pretrain the GRU-branch with the AMASS dataset [36] with synthesized IMUs as inputs as well as fitted joint kinematics as labels using Nimblephysics engine [62]. AMASS dataset is a large-scale motion capture dataset that include accurate kinematic labels of various activities. By pretraining on this dataset, the model can obtain a more robust kinematic estimation capability. To make the IMU synthesis closer to the real-world sensor readings, we randomize the IMU positions and add Gaussian noises to the synthesized IMUs. After the pretraining, we finetune the GRU-branch to better let the model capture the patterns in high dynamic conditions. Then the module is trained together with the other modules for the final KAM prediction. The workflow of this feature extraction module is shown in Fig. 5.

4.1.2 Body Segment Parameters Prediction. As mentioned in Sec. 3, besides kinematics information, we need body inertial information for better fulfillment of the inverse dynamics procedure. BSPs, namely segment mass, center of mass (CoM), and moment of inertia, are good target features that we can focus on to obtain body inertial information. Accurate obtaining of these parameters requires sophisticated anthropometric measurements like

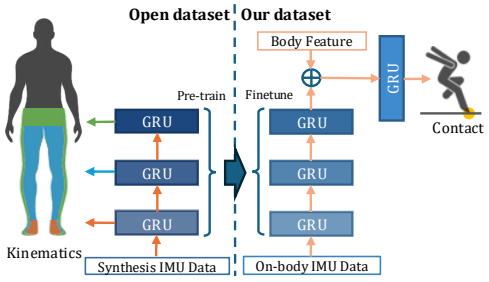


Fig. 5. Kinematics Feature Extraction. Small arrows demonstrate the data flow in forward operation. Middle arrow represents the pre-train and finetune workflow.

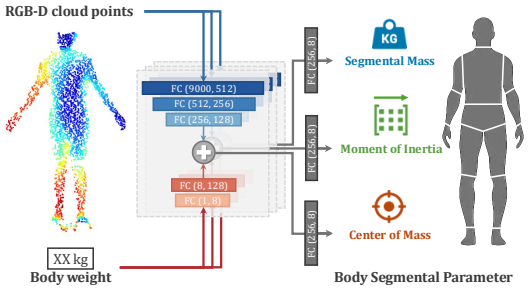


Fig. 6. BSP Estimation. Arrows demonstrate the data flow in forward operation.

CT or MRI [44], which are quite time-consuming and inconvenient for daily training usage. Recent papers [30] demonstrate that estimating these variables accurately from visual body scan is possible.

Another fact is that BSPs are essentially correlated with segment lengths, widths as well as volumes, therefore, it is possible to estimate BSPs if we can obtain accurate shape and volume measurement of each segment part given body weight is known. Inspired by this, we propose to use a one-time RGB-D registration method for capturing BSPs as shown in Fig. 6, making the measurement easy and practical in a daily training scenario. We utilize the point clouds generated by the RGB-D camera to estimate BSPs, which contain geometric measurement information in 3D space, as well as the total body mass for the regression of BSPs.

However, constructing the mapping from a point cloud of the whole body to the BSPs of each segment part requires prior knowledge of joint and segment positions. However, existing methods to extract joint positions will incur potential errors, leading to imprecise BSPs estimation. To avoid such error accumulation, we propose an end-to-end approach here with data-driven models. Specifically, we utilize a 3-layer multi-layer perceptron (MLP) to capture segmental information automatically without explicit joint positioning. To provide the posture prior, we only need to sort all coordinate points within the point clouds along with three axes. Note that even if we only construct half of the point clouds of the whole body, our method can still restore the whole body metric with the assumption that the front half of the body shares similar features as the back one. The evaluation results in Fig. 11 demonstrate the effectiveness of this approach. By automatically analyzing point cloud data using MLP, we can obtain features regarding the body shape. Additionally, we design another MLP to project the body weight into feature space and fuse with the body shape for BSPs estimation. And to extract inertial features from BSPs supervisions for further KAM regression, we utilize another MLP before the next stage. With such a design, we are able to obtain body inertial features in a simple setup and without training area constraints.

4.2 Physics-aware KAM Modeling

With sufficient biomechanical information retrieved, next we will illustrate our physics-aware KAM modeling framework.

4.2.1 Inverse Dynamics-based Multi-task Learning. In order to achieve an accurate KAM modeling based on the information we've obtained, we design a multi-task learning framework with respect to the principle of inverse dynamics.

From the equation we introduced in Sec. 3, we need to know about GRF term $J_s^T \lambda$ and Inertia Matrix $M(q)$ to get joint moments:

$$\tau = M(q)\ddot{q} + h(q, \dot{q}) - J_s^T \lambda \tag{3}$$

here we combine gravitational and Coriolis effect term into one term $h(q, \dot{q})$. Note that this term is relatively stable due to the near fixed impact of earth rotation and mass on the body.

The Mass Matrix $M(q)$ represents the inertial effect of the multi-body system of humans, which can be determined by each segment's BSP as well as the joint-space kinematics. Though determined, the transformation from BSP and joint motion to mass matrix is rather complex and requires prior knowledge of the human musculoskeletal system. Moreover, slight errors in joint motion estimation may result in erroneous mass matrices in deterministic calculation systems, leading to a negative impact accumulation on KAM prediction. Therefore, instead of using the deterministic approach, we use a data-driven approach and design an auxiliary task on regressing Mass Matrix from motion features as well as BSP features.

The GRF term $J_s^T \lambda$ represents the external force exerted on the multi-body system. The λ gives the amplitude of GRF while the Jacobian J_s determines how the contact point force will impact the whole system, which is a matrix whose shape is $3N \times M$ that maps GRF into the joint space (N is the number of contact points and M is the number of DoFs). Though there is no deterministic way to calculate λ from motion, many recent works have demonstrated a strong correlation between human motion and GRF amplitude, and the potential to infer GRF amplitude from IMU sensors using data-driven approaches [25, 48, 56]. Moreover, from GRF's dimension of the physical quantity, we can infer that it is not only determined by motion, but also by masses encoded in BSP features. We hence propose an auxiliary task to regress GRF amplitude from motion and BSP features. In terms of the Jacobian Matrix, it can be determined based on the contact point locations as well as joint-space kinematics of the current state. From a similar reasoning as the Mass Matrix regression, we choose not to use a direct computational approach to calculate the Jacobian due to potential errors in kinematic and contact point estimation. And note that the dimension of Jacobian Matrix $J_s \in \mathbb{R}^{3N \times M}$ is rather large, leading to difficulty in regression. Instead we use a surrogate variable, Center of Pressure (CoP), as an auxiliary task, which contains the essential contact points related information encoded in the Jacobian matrix with fewer dimensions for a better multi-task performance. Specifically, CoP is the point where the GRF on the body can be considered to act as a single force. By determining both CoP and GRF value as well as combining the body kinematic status, we are able to model the whole GRF term's impact, including both Jacobian matrix and GRF amplitude. We utilize both motion and BSP features for the CoP auxiliary task as well since the contact position information is determined by joint space kinematics and body skeletal information.

With the above-introduced auxiliary tasks, the model should be able to capture essential mapping information for KAM estimation. The corresponding multi-task learning design is demonstrated in Fig. 4. The rationale for such ID-based multi-task learning framework is that the shared representation it learned from these biomechanically-related auxiliary tasks can potentially reveal the biomechanical structure of human body and therefore help estimate KAM significantly. For the modeling of these three tasks, we design two GRU-based blocks to capture the related temporal features from the inputs, given its ability to model non-linear and complex correlations. Each GRU block contains one-layer GRU and one linear layer for feature extraction and dimensionality reduction. Since GRF and CoP describe the same term from different perspectives (value and direction), we choose to use the feature embedding extracted from the same GRU block. Each branch receives the concatenation of BSP and motion features extracted from BIR module and proceeds to further analysis. The final regressions are all performed using 2-layer MLP in these auxiliary tasks. Note that for different regression tasks, we use different feature extractors to avoid feature-wise cross-talk, since the features required for GRF term and Mass Matrix term are rather different.

4.2.2 BSP and Kinematics-guided Mixture-of-Experts KAM Regression. On top of the learning of three auxiliary tasks based on ID principles, we can next construct our KAM regression model. Given all the essential information and features we have provided, the regression is still difficult.

There are two main obstacles that limit the regression performance: Firstly, KAM pattern varies a lot due to the nature of highly dynamic movements in ACL injury prevention training. A subtle movement pattern change, like a slight knee adduction, will lead to a great KAM fluctuation, increasing the difficulty in balancing between providing enough fitting capacity and preventing potential overfitting. Moreover, though in principle BSP information is complementary to the motion information, the near-static characteristics of BSP may barely lead to an improvement in KAM estimation. For different trials of the same subject, the BSP features are the same since there is only one depth input. It is quite challenging to maximize BSP features' utilization in a neural network model.

Towards the first obstacle, we divide into the changing patterns of KAM and find that the variation of knee adduction angle contributes to a large portion of variance in KAM peaks and patterns due to the biomechanical correlation. Research demonstrates that knee adduction angle has a strong correlation with KAM patterns [52]. The rationale is that, KAM is enlarged when the direction of ground reaction force is not parallel with the tibia, and such non-parallelism is often caused by adduction movement of knee joint, which directly leads to adduction angle change. Therefore, we are inspired to let the model focus on the changing pattern of knee adduction angle. Though we have kinematics-guided feature extraction in the BIR module, the features are more focused on the general body posture instead of certain joint angle. We therefore design additional prediction branches on left and right knee adduction angles to make the model's attention on the adduction pattern of knee joints.

Moreover, since the KAM patterns in ACL injury prevention are rather unstable and varying compared to normal activities like walking, the mapping from motion to KAM curve is rather complicated, which requires deep models. However, heavy overfitting will happen given the mismatch between the large model capacity and limited data size. To better utilize the prior that knee adduction angles are strongly correlated with KAM patterns, we propose to utilize a Mixture-of-Experts (MoE) design in the KAM regression, to capture the complex mapping from different perspectives guided by priors while preventing potential overfitting. In detail, there are six MLP-based experts, each is a 3-layer MLP. The final output of the MoE module is weighing of all the experts based on the switch, which is a 2-layer MLP with softmax. Instead of directly concatenating the feature learned by knee adduction angle regression, we utilize its output as part of the input of the MoE switch, so that the MoE can automatically allocate experts to learn different KAM patterns by knee adduction angle prior information.

Towards the second obstacle, to maximize BSP's utilization, we review the BSP's impact on knee adduction moment. In principle, BSP can reflect the biomechanical structure of a human body, determining if a subject is tall and thin, or short and fat. And based on the biomechanical structure of subjects, their movement patterns can be clustered to some extent, as human motion patterns are largely determined by the biomechanical structure. Therefore, we propose to make a better BSP utilization through two designs. We add feature-wise linear modulation of BSPs on top of the motion feature instead of direct concatenation to make a better feature fusion.[46] The rationale is that comparing to full attention of the feature, such technique can fuse features with less parameters, preventing the network from potential overfitting. We also use BSPs as part of the input to our MoE switch. The rationale for such BSP-enhanced MoE design is that KAM is naturally correlated with body parameters like weight and height, which may have pattern clustering phenomenon among subjects with similar biomechanical structures. MoE can capture such clustering by different experts, and therefore provide better generalizability.

Apart from the above-mentioned BSP and knee adduction angle-guided MoE design, we further propose to enhance the final KAM regression with cross-attention. Note that we have mentioned that we split the left and right lower limbs' motion features, in KAM regression we also split the right and left leg to prevent potential cross-talk inside the regressor, since the range of KAM of the left leg and right leg might differ due to the control ability of dominant and non-dominant leg. However, we still need to fuse the information from both legs as well as all the features extracted from the ID-guided multi-tasking. The rationale is that, when the knee adduction moment is abnormal for one leg, the other leg might be in an unstable posture for dynamic body mass center control during sports training [37]. To optimize the feature fusion, we propose to use a multi-head cross attention

module on top of the concatenated features. We utilize the corresponding lower limb's feature as key and value while the other leg's feature as query to emphasize the potentially important temporal positions in the other leg's perspective. We also design a residual connection on top of the attention module to preserve as much information in the original features.

To better train our neural network, we design different loss functions atop our PKM module. Specifically, the following are the loss functions:

$$\mathcal{L}_{main}(y_i, f(x_i), \alpha_i) = \alpha_i * \begin{cases} 0.5 * \|y_i - f(x_i)\|^2 & , \|y_i - f(x_i)\| \leq \delta, \\ \delta * (\|y_i - f(x_i)\| - 0.5 * \delta) & , \|y_i - f(x_i)\| > \delta \end{cases} \tag{4}$$

We adopt the weighted smooth quadratic loss here as our main regression function for KAM regression and all the other auxiliary tasks. Note that y_i here is the ground truth label and $f(x_i)$ is our prediction. δ controls the quadratic region of the loss, and α_i is the weight. The rationale for using this loss is that the convergence with general MSE loss may be affected by pattern outliers, leading to inferior performance in most cases. In our KAM regression, the α_i value is set to be the absolute value of ground truth, i.e. $\|y_i\|$, to emphasize the high-risk region with higher absolute KAM values which we should pay more attention to. The δ value we used in KAM regression is 0.5. In other auxiliary tasks, the α_i and δ values are 1.0.

In addition to this, we further utilize frequency loss function in KAM regression:

$$\mathcal{L}_{freq}(y_i, f(x_i)) = \|\mathrm{FFT}(y_i) - \mathrm{FFT}(f(x_i))\| \tag{5}$$

Here the FFT function means the real fast Fourier transform on the given time series. The motivation for using this loss is that, smooth quadratic loss on the time domain focuses on lower frequency components of the KAM curve in prior, and potentially ignores the higher frequency part. For instance, there are peaks or the points with sudden directional changes on the KAM curves, which are important for regression error reduction. We guide the model to emphasize these points by using this frequency loss. Apart from this, in order to avoid experts collapsing in our MoE architecture, we use entropy loss on the switch:

$$\mathcal{L}_{entropy}(\mathrm{switch}_i) = -\sum_{i=1}^{N} \mathrm{switch}_i \cdot \log(\mathrm{switch}_i) \tag{6}$$

so that the switch will consider all the experts' outputs in a near-equal manner. The final loss function will then be:

$$\begin{aligned} \mathcal{L} = \sum_{i \in (L,R)} & (\alpha \mathcal{L}_{main}(\mathrm{KAM}_i) + \alpha \mathcal{L}_{freq}(\mathrm{KAM}_i) + \beta \mathcal{L}_{main}(\text{knee adduction angle}_i) \\ & + \mathcal{L}_{entropy}(\mathrm{switch}_i)) + \sum_{i=0}^{3} \mathcal{L}_{main}(\mathrm{Out}_{aux_i}) \end{aligned} \tag{7}$$

where the α is set to 5 and the β is 3 to increase the weight on main KAM as well as knee adduction angle regression, and Out_{aux_i}refers to the output of each auxiliary task prediction.

5 EVALUATION

In this section, we will introduce our evaluation, including both study setup and results.

5.1 Study Setup

We first illustrate our enrollment and data collection details, followed by ground truth and training details.

Fig. 7. High-risk Tasks Illustration. AD stands for knee adduction and AB stands for knee abduction.

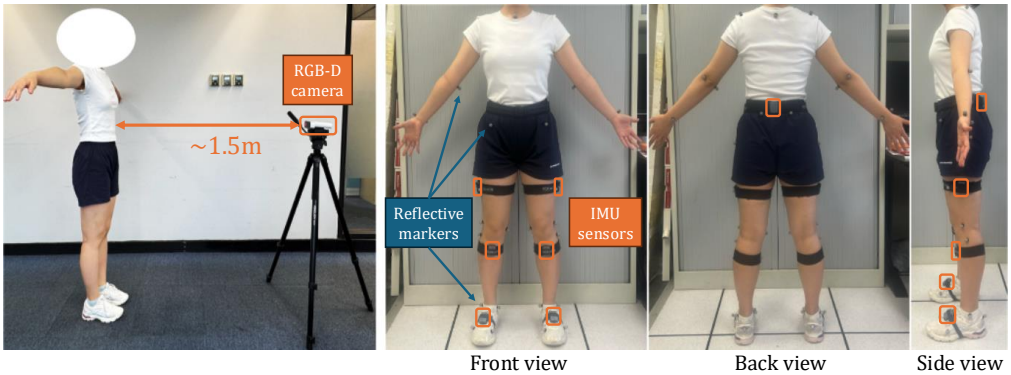


Fig. 8. Setup Illustration.

Table 1. Participants Statistics.

Stats	Athletes	Non-athletes
Population	10	9
Age (years)	24.7 (2.0)	24.2 (1.1)
Height (cm)	169.8 (7.9)	174.8 (10.2)
Weight (kg)	61.2 (10.6)	63.8 (10.6)
Body Mass Index (kg/m^2)	21.2 (3.0)	20.9 (2.8)
Gender	Female: 5 Male: 5	Female: 3 Male: 6
Tegner Score	6.1 (0.3)	4.0 (1.2)

Statistic representation: Mean (SD).

5.1.1 Participants. We recruited 19 participants, including 10 athletes and 9 non-athlete subjects on the basis of research protocols approved by the institutional review board. Written informed consent is obtained from all subjects. Participants who score over 5 in the Tegner scale are considered as athletes with respect to the standard in biomechanics research [9, 16]. All the participants have no reported knee pain symptoms or injury history. We briefly present participant-related statistics in Table 1.

5.1.2 Devices. We utilize Microsoft Azure Kinect [38] as our RGB-D camera, and seven Noraxon Ultium Motion IMUs [2]. In terms of the biomechanical system, we utilize Vicon motion capture system [3] as well as Bertec force plate [1] to capture the motion and GRF data. We synchronize IMUs and the biomechanical system at a sampling rate of 200Hz, while the GRF data is captured at a sampling rate of 1000Hz to capture detailed kinetic information.

5.1.3 Procedure. Before each data collection session, the participants receive a comprehensive overview of this study and a detailed instruction. Then they will go through the one-time body scan registration, wear the sensors on their body with the help of the researcher. After that the participants will be taught the high-dynamic tasks first and then perform them one by one.

- **One-time RGB-D Registration.** The participants are asked to stand still in a T-pose in front of the RGB-D camera. The distance between the user and the camera is around 1.5 meters. The reason is that the field-of-view of Microsoft Azure Kinect in around 1.5 meters will cover the whole bodies of all our subjects. The subjects are wearing tight T-shirts and sports shorts for more accurate BSP estimation, which are also common practice in sports training settings.
- **On-body Sensors Setup.** For the reflective marker setup for the motion capture system, we place 34 markers on the participants' body, mostly based on the Rajagopal human model's sample marker placement [47]. In terms of the inertial sensor setup, we place seven sensors on the participants' lower limb body parts via elastic bands. We set one IMU on each lower limb segment and one on the pelvis, and adjust the position on thighs for soft-tissue artifacts reduction [63]. In detail, we set: pelvis IMU at the center of the waist; thigh IMUs at the middle of iliotibial bands, which are the centers of thighs in the side view; shank IMUs at the position that is just below tibial tuberosities, and facing the front side of the shanks; foot IMUs at the arches of the feet. Such a lower limb IMU setup set is a common practice for on-field monitoring of sports training with respect to the recommended Noraxon IMU setup. The sensor placing illustration is shown in Fig. 8.
- **High-dynamic Tasks.** Four typical high-dynamic tasks will be performed by participants six times each, including **double-leg jump-landing (DJ), single-leg jump-landing (SJ), cutting (CT) and pivoting (PT)** with respect to common ACL injury prevention training tasks [24, 55]. The participants are asked to perform all tasks on the force plates. We additionally ask participants to perform two trials with more knee abduction (AB) and other two trials with more knee adduction (AD) to involve more movement variations that will happen in real-world scenarios. Due to the difficulty in controlling the knee position during PT, we only ask the participants to make such movement variations during the other three tasks. The demonstration of the tasks is shown in Fig. 7.

5.1.4 Ground Truth. Based on the captured motion trajectories and GRF, we calculate the KAM with OpenSim's inverse dynamics tool [15], which is a widely accepted ground truth standard in biomechanics. In detail, we use the Rajagopal model as our base human model and scale the model based on the static measurement test of each participant. And for the auxiliary tasks, the ground truth value of mass matrix as well as BSPs are also obtained from OpenSim. The GRF and CoP ground truths are obtained directly from the force plate measurement.

5.1.5 Training Setting. We implement our neural network models with Python 3.10 and PyTorch 2.5 [43]. All the evaluation experiments are conducted on NVIDIA GeForce RTX 4090 GPUs. We utilize the leave-one-subject-out scheme for all of our evaluations, with the test subject's data unseen during training. We set the initial learning rate to be 1e-3 for the training of joint kinematic features' GRU branch as well as the BSP modeling before the overall training. The rest training procedure is performed also with an initial learning rate at 1e-3 and a cosine annealing warm restarts scheduling scheme, while the finetuning learning rate of GRU branch is 5e-4. We use AdamW [32] as our main optimizer. The sequence lengths of inputs vary based on different trial lengths. The details of the model parameters are provided in the open-sourced code repository.

5.1.6 Benchmark and Baseline. To demonstrate the performance of ACLGuard, we compare with the reported results in state-of-the-art Markerless MoCap method [58] as well as IMU-based method [54] in high-risk scenarios. To the best of our knowledge, there is no related work that combines IMU sensing as well as one-time RGB-D registration for KAM estimation in this scenario, so we refer to the state-of-the-art IMU-based related work and adopt a similar MLP structure as our baseline to evaluate on our dataset. In detail, the baseline model is a

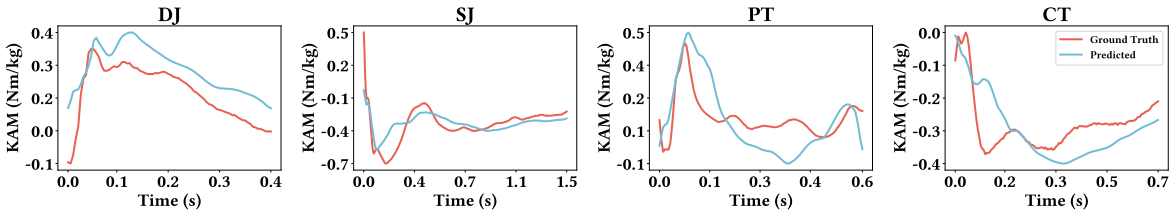


Fig. 9. Sample Prediction and Ground Truth.

Table 2. Overall Performance.

Benchmark	**KAM RMSE (Nm/kg)↓**
Markerless Mocap [58]	0.16
ACLGuard	**0.18**
IMU-based [54]	0.53

three-layer MLP, with hidden sizes of 100 and 20 and hyperbolic tangent function as activation with respect to the paper.

In addition to the above work that is most relevant to our scenario and setup, we additionally compare ACLGuard with two models presented in other state-of-the-art work that also leverage IMUs to estimate human kinetics metrics like joint moment and ground reaction force. We refer to [56] which designs a 6-layer Transformer-based architecture for ground reaction force estimation, and [39] which proposes a temporal CNN (TCN) design for joint moment estimation as our baselines as well.

5.1.7 Evaluation Metric. To evaluate the accuracy of KAM estimation of ACLGuard, we utilize two main metrics, the root mean square error (RMSE) and normalized root mean square error (NRMSE). These two metrics are commonly used in related papers [54, 58, 63]. The calculation formula for RMSE is as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

where y_i is the ground truth value and $\hat{y}_i$ is the prediction value. The unit of KAM is $N \cdot m/kg$ since we normalize the joint moment value by the subject's body weight, which is standard practice in joint moment prediction.

As the absolute value of KAM is non-intuitive to be understood, we additionally utilize NRMSE in our evaluation. The calculation formula of NRMSE is:

$$NRMSE = \frac{RMSE}{\max(y) - \min(y)}$$

where y is the ground truth value of KAM. NRMSE presents a view of percentage error in KAM regression task. As the KAM range will vary from different people, we utilize subject-wise normalization in this NRMSE calculation. In the overall and ablation evaluation, we report both RMSE and NRMSE for comparison, while in robustness and subject-wise study, we utilize NRMSE only as it can compare subject-wise performance in a more fair manner.

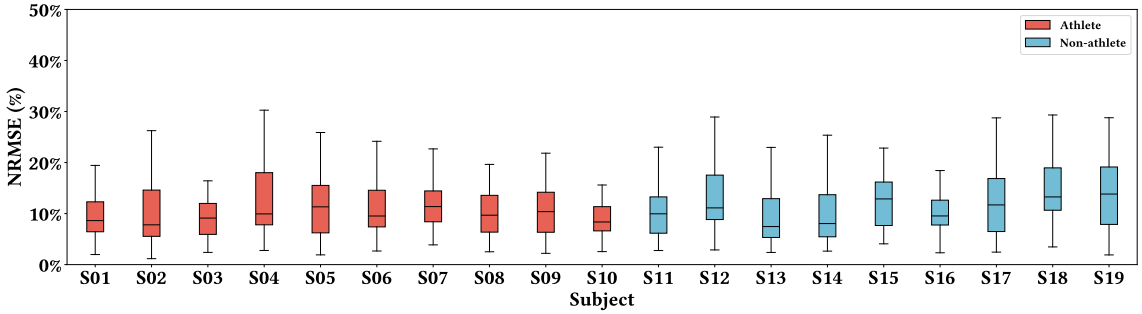


Fig. 10. KAM Estimation Performance of Each Subject. The athletes and non-athletes are shown in red and blue correspondingly. The middle line show the median NRMSE of each subject.

5.2 Overall Performance

5.2.1 Baseline Comparison. From the Table. 2, we can see that ACLGuard can achieve a comparable error level with the existing state-of-the-art Markerless Mocap method in high-risk conditions. In addition, it greatly outperforms existing IMU-based solutions by reducing over half of the error level. This result demonstrates the effectiveness of our system, as we are able to achieve an acceptable performance without the limitations of occlusion and out-of-view in vision methods.

Meanwhile, Table. ?? shows that ACLGuard performs better than all the baseline methods evaluated on our dataset. Notably, ACLGuard's performance surpasses the best performing Transformer architecture among all the baselines by reducing more than 25% mean RMSE and decreasing NRMSE level to our 10% as well as achieving smaller inter-subject standard deviation, showing the benefit of our designs. To give a more intuitive view of our prediction performance, we provide sample prediction curves of each movement task conducted in Fig. 9, where we can see that our system can not only predict the overall curve changing well, but also estimate the peaks at an accurate level. Related works have discussed that around 10% NRMSE level is of a reasonably good level of accuracy in terms of joint moment estimation [31], which can show that our system's performance is acceptable for potential real-world usage.

5.2.2 Subject-wise Performance. In addition, we here present our system's performance on each individual subject. The results are shown in Fig. 10, which uses box-plots to demonstrate the trial-wise error distribution within each subject. The subjects are divided into two groups with respect to whether they meet the criterion of athlete or not based on the Tegner scale. To mitigate the impact of KAM value range difference among different subjects, we utilize NRMSE in this evaluation. From the results we can see that, our system can achieve a relatively consistent estimation performance across different subjects with all the subjects' mean NRMSE level is around 10%, while the intra-subject error rate variation is a bit high, which may be due to the varying movement tasks and large movement variations we introduced in our data collection protocol. Notably, the variation in athletes population is slightly smaller than the normal control group, while the mean error level is similar, which may be a result of athletes' stronger capability in controlling their movement in high-risk tasks while the non-athletes will have a larger variation in high-risk movement patterns. Overall speaking, the subject-wise performance shows the robustness of our system's performance among different testing subjects.

Table 3. Benchmark Comparison.

	KAM	
	RMSE (Nm/kg)↓	NRMSE (%)↓
ACLGuard	**0.176 (0.030)**	**11.5 (1.65)**
Transformer [56]	0.239 (0.035)	15.4 (2.09)
TCN [39]	0.261 (0.031)	17.1 (2.55)
MLP [54]	0.293 (0.031)	19.1 (2.25)

Stats representation: Mean (SD). Best results are in **Bold**.

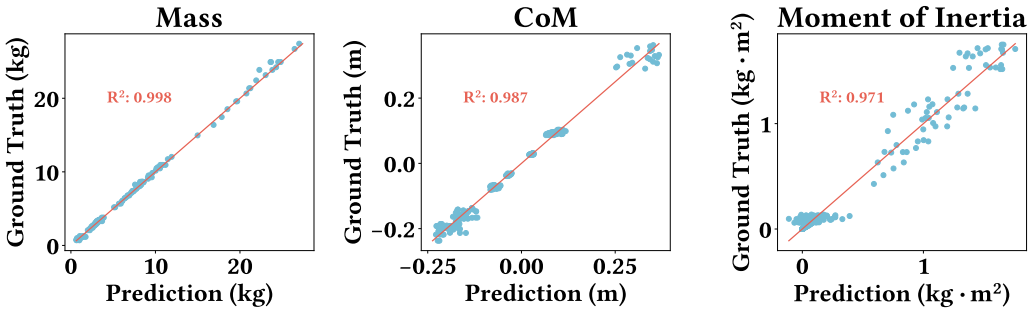


Fig. 11. Performance of BSP Estimation.

5.3 Modular Evaluation

Next, we evaluate the effectiveness and contribution of system modules we've designed. We illustrate the effectiveness of our estimation on BSPs utilizing point clouds captured by RGB-D camera, and presents the ablation studies of all the designed modules.

5.3.1 Performance of BSP Estimation. The demonstration of RGB-D based BSPs estimation accuracy is demonstrated in Fig. 11. We present three separate scatter figures to show the regression accuracy of three parameters of BSPs: segmental mass, segmental CoM and moment of inertia. The red line is provided as a reference line of predictions and ground truths are the same. We collectively demonstrate the performance of the segments of all subjects in these figures. From the results we can see that, all the BSPs are predicted with a strong correlation against the ground truth value. For segmental mass estimation, the ground truth from OpenSim is a linear scaling of the total mass measured in principle, which leads to a strong linear correlation as we also have total mass as well into our BSP estimation module. In terms of CoM and Moment of Inertia, the mappings in OpenSim from the marker and mass measurement to them are more complicated, which involves locating bony landmarks and conducting distance measures in the musculoskeletal model, leading to a larger variability in the error rate. Nevertheless, the overall prediction performance is still good for all segments and all subjects, demonstrating the effectiveness of the BSP module.

5.3.2 Benefit of BSP Module. Next we will explain the benefit of BSP module. We perform an ablation study by replacing genuine BSP inputs with pseudo BSPs of the subjects. The result is shown in Case 2 of Table. ??. We can see that by comparing Case 1 and 2, both the RMSE and NRMSE increased, demonstrating the effectiveness

Table 4. Ablation Studies

Case	BSP	KFE	MTL	MoE	KAM RMSE (Nm/kg) (↓)	KAM NRMSE (%) (↓)
1	✓	✓	✓	✓	**0.176 (0.030)**	**11.5 (1.65)**
2	✕	✓	✓	✓	0.187 (0.028)	12.2 (2.03)
3	✓	✓	✓	✕	0.189 (0.032)	12.3 (1.91)
4	✓	✓	✕	✕	0.270 (0.063)	17.8 (4.33)
5	✕	✕	✓	✓	0.213 (0.036)	13.9 (2.17)

↓ indicates a lower error.
BSP: Body Segment Parameters Prediction.
KFE: Kinematic Feature Extraction.
MTL: Inverse Dynamics-based Multi-task Learning.
MoE: BSP and Kinemaitcs-guided Mixture-of-Experts.

of the BSP estimation module. In addition, for NRMSE part, the standard deviation level also increased a bit, showing that BSP module can help stabilize the overall prediction performance. It is worth mentioning that, though erroneous BSP information is provided in this ablation study, the system might inferred related physical constraints through inverse dynamics-based multi-task learning procedure.

5.3.3 Benefit of KFE Module. By comparing Case 2 and 5, we are able to decode the effectiveness of the KFE, or kinematic feature extraction module. For the ablation study conducted in Case 5, we remove the feature extraction module for both kinematic as well as BSP part, only keeping the raw IMU and one-time RGB-D registration data as input to ensure the information is the same. We can see that comparing to Case 2, Case 5 has an increase in both mean RMSE and NRMSE. Such benefit comes with the motion feature extraction ability provided by both pretrained GRU-based and ResUNet-based architecture.

5.3.4 Benefit of MoE Module. In addition, we evaluated the benefit of BSP and kinematics-guided MoE module by comparing Case 1 and 3. In Case 3 we replace the MoE module with a simple MLP for ablation study. We can see that both the mean RMSE and NRMSE increased, showing the gating scheme and MoE design can effectively reduce the overall estimation error level.

5.3.5 Benefit of MTL Module. We also validated the benefit of MTL module by comparing Case 3 and Case 4. In Case 4 we remove the whole physics-aware KAM modeling module. We can see significant performance degradation after removing the MTL module, in terms of both mean error level and variation level. This can be explained by the rationale that ID calculation will require parameters including GRF, CoP and Mass Matrix. By predicting these auxiliary tasks, the backbone of neural network will learn a better hidden space to model kinetic metrics from motion and body features. With such a MTL design, the system are able to make better use of captured motion signals and body measurements towards the KAM estimation.

5.4 Multi-task Learning Dynamics

In this section, we will provide training dynamics and performance of the inverse dynamics-based multi-task learning module of our system.
As the nature of multi-task learning will involve potential training instability, we here demonstrate the mean evaluation loss curve on all test subjects in Fig. 13. We can see that in all three auxiliary tasks, the evaluation loss

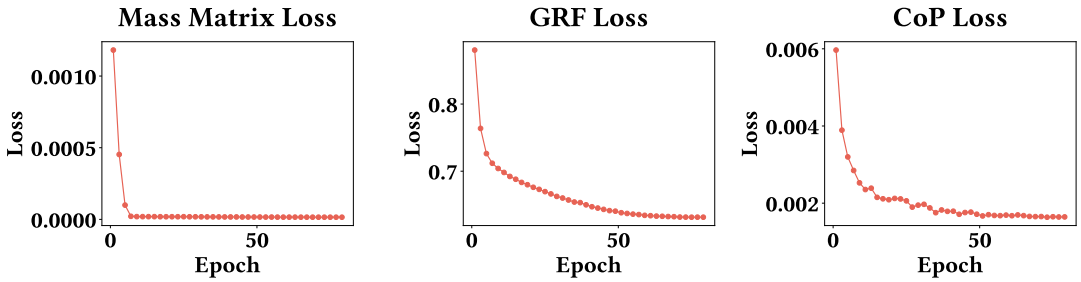


Fig. 12. Multi-task Learning Dynamics.

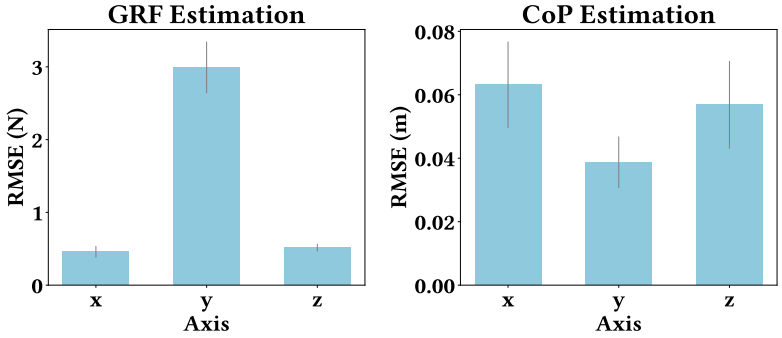


Fig. 13. Multi-task Learning Accuracy. y-axis corresponds to the axis of gravity.

all drop smoothly as the training converges. The mass matrix regression converges most quickly while the GRF converges slower. However, after the training goes to around 60 epochs, all three auxiliary tasks converges. This demonstrates the effectiveness of our training procedure in these tasks.
Moreover, we here provide the estimation accuracy of these auxiliary tasks in the final model. As the mass matrix is high dimensional ($N \times N$, where N is the DoF of the whole human body) and its value is not intuitive enough, we here only provide the prediction result of GRF and CoP. We can see that, the largest GRF estimation error is on the y-axis (the axis of gravity.), which is only around 3N, and the CoP estimation errors of all three axes are around 5cm. These error levels are quite small, especially in highly dynamics scenarios, where the value of these parameters may change drastically, showing the effectiveness and benefits of our system's MTL design.

5.5 Robustness Analysis

To investigate the system's effectiveness under different subject-wise or activity-wise conditions, we conduct a robustness study with respect to five impact factors in real-world scenarios, including types and variations of movements, subject demographic characteristics and sensor setups.

5.5.1 Impact of Movement Types. One major impact we want to investigate is that of different movement types. During our data collection, we ask all subjects to perform four major types of high-risk tasks that commonly practiced in ACL injury prevention training, namely double-leg jumping (DJ), single-leg jumping (SJ), cutting (CT) and pivoting (PT). Notably, from Fig. 14a and 14b, we can see a relatively smaller mean error and standard deviations for DJ and CT trials, while larger mean error level as well as variations in SJ and PT. This is due to

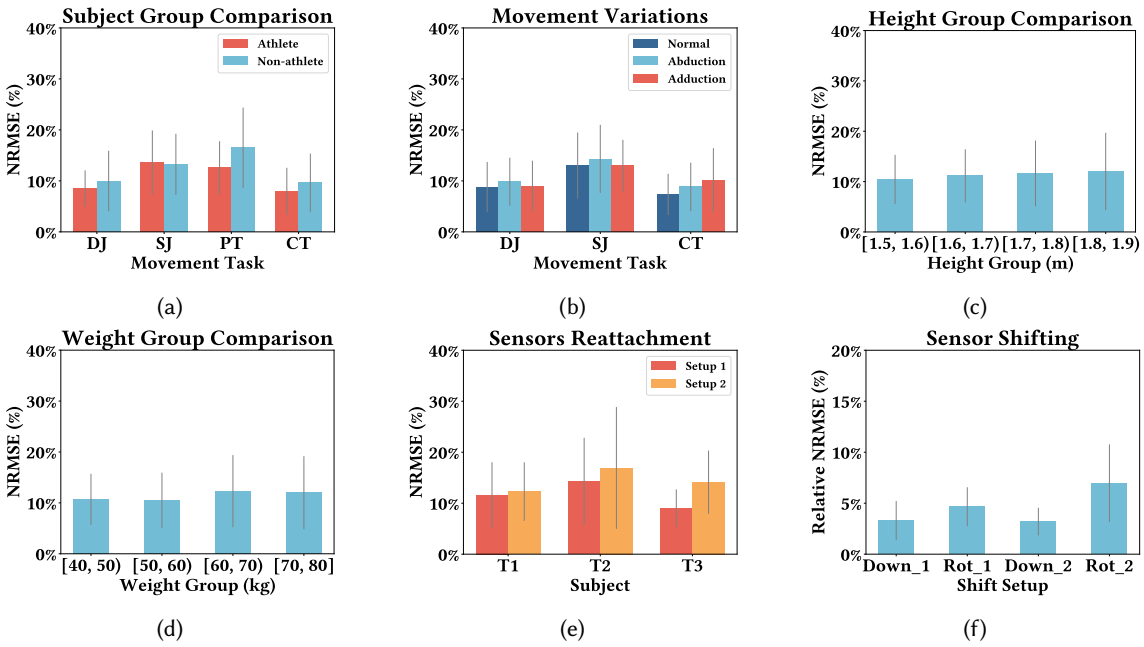
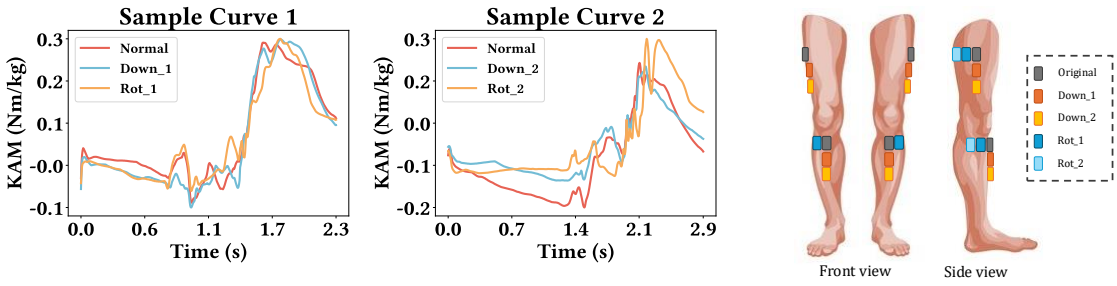


Fig. 14. Robustness Study. (a) and (b) presents KAM estimation performance differences with respect to different subject groups, different movement variations and tasks. (c) and (d) shows estimation performance differences among different height and weight groups. (e) and (f) demonstrates system's robustness under sensor reattaching and shifting conditions.

the fact that SJ and PT can be comparatively more dynamic and of higher landing impact than the other ones, leading to more variations and larger value ranges compared to DJ and CT. However, the mean error level of SJ and PT are not significantly higher compared to the other ones, with around 10% NRMSE. Given the higher value range of SJ and PT, we think the system's overall performance is still robust against different movement types.

5.5.2 Impact of Subject Groups. Next we consider the impact of subject groups. We divide all the subjects into athlete group and non-athlete group with respect to their Tegner scale, and report their trial-wise mean NRMSE and standard deviations with respect to their movement tasks in Fig. 14a. We can see that apart from the previously discussed performance differences in movement types, there is a slight mean error level increase but not significant in non-athlete group comparing to athlete group in three tasks, which echos our previous discussion in Sec. 5.2.2. But overall, both the error means and variations are of the same level, demonstrating the robustness of our system against different types of users.

5.5.3 Impact of Movement Variations. To make our evaluation closer to the real-world prevention training scenario, we asked the participants to perform tasks with three different types of movement variations as described in Sec. 5.1: land with natural knee position, land with knee abduction and land with knee adduction. Due to the task difficulties, we only ask subjects to perform such movement variations in DJ, SJ and CT tasks, and report task-wise and movement variation-wise performance in Fig. 14b. From the figure, we can see that the error rate is relatively same in both adduction and abduction poses for DJ and SJ tasks. While for CT task, abduction and especially adduction pose will lead to some increment in error rate. The reason for this might be due to potential knee instabilities in such movement variations during CT. And in SJ task, the mean error



(a) Sample Curves Comparison.

(b) Setup Illustration.

Fig. 15. Impact of Sensor Shifting.

rate is generally higher due to the inherent task difficulty as the subject need to keep balance, which covers the performance differences due to movement variations. In DJ task, the overall body movement is more stable, which leads to a more consistent error rate under different posture variations. Nevertheless, even for the adduction pose in SJ task, the error rate is still below 15%, demonstrating the system's robustness against movement variations.

5.5.4 Impact of Subject Body Measures. We move on to discuss another two potential subject-wise factors, which are the heights and the weights of subjects. We here analyze the impact of these two factors and see if they will impact the robustness. The results are given in Fig. 14c and 14d. We can see that the performances of our system are around the same level for different height and weight groups, while the larger heights and weights seem to give slightly bigger mean error rates and standard deviations. The reason behind this might be that those subjects will have a larger impact load onto the ground in highly dynamic tasks, which leads to some error increment. However, the error levels are still of the same level and acceptable in general.

5.5.5 Impact of Sensors Reattachment. We further conduct one robustness study, which involves three subjects performing the same experiments in two days, with all the IMU sensors reattached to certain body parts. Since we cannot ensure the placement of wearable sensors to be exactly the same across different trainings, we use such a performance comparison to demonstrate our system's robustness to the setup differences caused by sensor reattachment. We don't include all data of that testing subject from both days when evaluating reattachments. From the result we can see that, there are some but no significant changes in the error levels across setups for the first two subjects, and only a slight increase of NRMSE of T2 and T3 in second day's setup, but all six trials' mean error rate is in an good region around 10%, which is acceptable and demonstrates the system's robustness on sensor reattaching setups.

5.5.6 Impact of Sensors Shifting in Real-world Settings. To further investigate the impact of real-world deployment issues, we here demonstrate another robustness study regarding the influence of sensor placement shifting on our system, which is shown in Fig. 14f. We consider two main cases of sensor shifting: *downward slipping* (Down_1 and Down_2 in the Figure), which is the most common shifting case that happen after multiple high dynamic movements like jumping and cutting; *rotary shifting* (Rot_1 and Rot_2 in the Figure), which will appear when there are heavy soft-tissue artifacts during training. Number 1 and 2 in the label refer to the extent of shifting, in which setup 2 is more deviated from the normal setup position. The schematic diagram of the wearing setups is provided in Fig. 15b. Since the shifting will mostly happen in thigh and shank, in this experiment the setup changes are performed in thigh and shank of both legs.

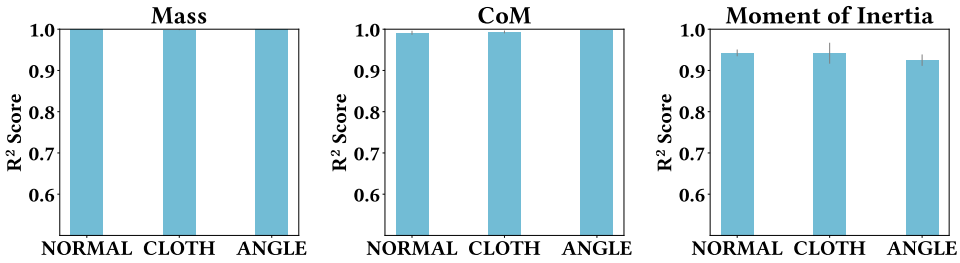


Fig. 16. Impact of Body Variability on BSP Estimation.

Table 5. System Latency.

Device	BIR Module	PKM Module	Total
CPU	138.3 ms	37.2 ms	175.5 ms
GPU	34.0 ms	3.9 ms	37.9 ms

In this evaluation we collect two subjects' data, and don't include all data of that testing subject from the training set when evaluating. Since there will be no ground truth motion capture system in the out-of-the-lab settings, we here provide an alternative strategy to validate the system's robustness. We simultaneously set and record IMUs in both normal and shifting setups, and compare the prediction results of both setups to demonstrate the model's robustness against different IMU setups. The metrics provided here is relative NRMSE comparing to the prediction of normal setup to demonstrate the system's prediction consistency. We can see that for both the downward slipping cases, the relative NRMSE is only around 3%, demonstrating that the system is robust against the case that the IMU band slip down a bit. Comparatively, the estimation differences to normal setup of rotary cases are larger, showing that the relative attitude of IMU to the segment will indeed influence the system's performance to some extent. To demonstrate our system's robustness in a more intuitive manner, we here provide the sample KAM estimation curves in Fig. 15a.

5.5.7 Impact of Body Variability in RGB-D Initialization. We additionally explore the impact of body variation factors on RGB-D initialization and BSP estimation part. Since we utilize vision-based methods, there might be potential influence from clothing, shooting conditions and etc. We here investigate two main factors, which is the clothing change and shooting angle difference, by additionally collecting one-time RGB-D registration of two subjects that are involved in our main dataset but with different clothing or shooting angle setup. During main dataset collection, we ask the subjects to wear tight clothes for better BSP estimation. We here ask the test subjects to wear loose T-shirts for shooting. For evaluating the shooting angle difference, we shoot from the front-left side of the subject instead of the front side to see its impact. Fig. 16 demonstrates the mean R^2 Score and standard deviations of two subjects' BSP estimations. We can see that, for the Segmental Mass and CoM estimation, the impact of both clothing and shooting angles are small, while for moment of inertia, the shooting angle change degrades the BSP estimation while the impact of clothing change is also not that significant. This can be explained by the shooting angle change will largely alter the distribution of point clouds, leading to a relatively poor estimation performance. Nevertheless, the R^2 Score is still around 90%, showing the robustness of our BSP estimation module.

5.6 Model Inference Analysis

In real-world deployment scenario, the model will be run on the coach's laptop and showing the KAM curve after one training trial, normally several seconds, and the coach will accordingly give instructions to the athlete for adjusting their postures and movement patterns. Such a scenario requires the model to have a relatively low inference delay. Hence we here evaluate the inference latency of our system. The results are shown in Table. 5. Since the lengths of input sequences may vary depending on the trial duration, we here present the latency by normalizing the real delay into the number of milliseconds per one-second sample. We test the inference delay on AMD EPYC 7542 CPU and NVIDIA GeForce RTX 4090 GPU. We can see that the inference latencies on CPU and GPU are less than 200 ms per each one-second sample, indicating the potential of real system deployment in on-field scenario.

6 RELATED WORK

6.1 KAM Estimation in Normal Activities

Various solutions have been proposed to estimate KAM in normal daily activities. A vision-based system developed by Boswell et al. [8] uses 2D video analysis together with an automatic body landmark detection module to predict KAM during walking. The camera is placed right behind the users when they are walking on a treadmill. The system can achieve a Mean Absolute Error (MAE) of 0.53 and R^2 of 0.78. Tan et al. [57] propose to combine the smartphone camera and IMU to predict KAM in gait. The cameras are placed at the back and right side of the user, and eight IMU sensors are attached to their certain body parts. Its RMSE is 0.49 %BW·BH, where BW and BH are body weight and height. However, camera-based solutions are limited to certain areas, denying the deployment in wild scenarios. Therefore, wearable solutions based on IMU are proposed. Jung et al. [28] presents a system consisting of 9 IMU sensors to predict KAM during gait. A multi-model is designed to use data from IMU attached to lower limbs, whose RMSE is evaluated to be 6.84 Nm. Wang et al. [61] develop a real-time KAM feedback system. The system achieves an R^2 of 0.95 in KAM estimation during walking by leveraging two IMU sensors firmly affixed onto ankles. Another system introduced by Bernd et al. [54] leverages IMUs to estimate knee abduction moments during various locomotion tasks. For normal tasks of walking and moderate running, the RMSE could achieve 0.18 ± 0.06 and 0.37 ± 0.14 Nm/kg. The RMSE increases to 0.80 ± 0.46 Nm/kg when the running speed is faster, indicating that KAM estimation in rapid movement is much more challenging. A recent system proposed by Yang et al. [63] leverage both IMUs and sEMG to estimate knee adduction moment and muscle force during walking for gait retraining. The reported NRMSE can achieve 9.95% in terms of walking KAM. However, the sEMG sensor may have potential signal quality degradation issue in highly dynamic conditions [14], which may lead to performance degradation in injury prevention training scenario. All of the above-mentioned KAM estimation in normal activities may have issues in highly dynamic conditions due to the movement nature [54], therefore many methods have been proposed to achieve better estimation results in high-risk activities.

6.2 KAM Estimation in High-risk Tasks

Prior studies have also put efforts into KAM monitoring for high-risk tasks. Templin et al. [58] evaluated a novel 3D markerless motion capture system in kinematics and kinetics estimation. The researchers conducted drop vertical jump tasks with 127 athletes. The reported RMSE for knee abduction moments was 0.16 Nm/kg. Annamaria et al. [21] examined the ACL Quick Check system in lower limb joint kinematics and kinetics when performing double-leg squats. The system consists of cameras and plantar pressure insoles. Compared to the ground truth calculated by ID, the mean RMSE of KAM was 0.42 N/kg. The camera settings limit these solutions to certain areas. Moreover, both of them only validated the systems on one type of high-risk task. Wearable solutions based on IMUs have also been proposed, taking advantage of their mobility in free-area. Serena et al. [11] focus on knee joint moments estimation in vertical drop jump task. Three IMUs are attached to one

leg for estimation. The proposed system achieved a mean RMSE of 0.04 Nm/kg. Although the performance is acceptable, the system is only evaluated on one type of high-risk task performed by non-athlete subjects. The system developed by Bernd et al. [54] is also validated on high-risk tasks like pivoting and cutting. Compared to normal tasks, the performance degrades sharply on these two high-risk tasks. The RMSEs of predicted KAM value in pivoting and cutting are reported to be 0.62 ± 0.19 and 0.92 ± 0.54 Nm/kg respectively. In comparison, our system successfully estimates KAM in four typical high-risk tasks performed by both athletes and non-athlete subjects. Meanwhile, the proposed system could be potentially deployed for on-field ACL prevention training scenarios with acceptable performance.

7 DISCUSSION

Population Diversity. Though we have included our dataset with both athletes as well as normal subjects with respect to the Tegner scale, we did not study the impact of specialized sports types on the athletes. In addition, the age distribution of our dataset does not vary a lot, even if it is within the common age group of athletes. Since age and sports types will impact the ACL injury rate as well as their movement patterns, these factors may influence our system's performance. In the future, we will include a more diverse evaluation population to investigate the impact of these factors on our system.

Feedback Design. This paper demonstrates that ACLGuard can be used as a system to augment ACL injury prevention training by monitoring KAM. However, we haven't designed and evaluated the feedback scheme of KAM and its impact on the training effectiveness. But with a low complexity algorithm and accurate estimation performance, this system has the potential to deploy in real mobile devices and adapt to different biofeedback strategies like visual-based or audio-based. We will design and evaluate different feedback strategies on top of ACLGuard in the future to validate their effectiveness.

Wearing Comfort. In this paper, we set up the IMUs with elastic bands. In order to make them relatively stable to the body, the elastic bands need to be tightly tied. We conduct a short user feedback study on the participants. One of them feels a bit uncomfortable due to the tightness, while three athletes and one non-athlete subject say that they worry that such a system may affect their sports performance. In the future, we will try to re-design the form factor of the wearables into a smart fabric-based system, to capture joint motions and kinetics with a more user-friendly wearing experience.

Setup Requirement. In our system setup, the distance between the RGB-D camera and the human body is around 1.5 meters due to the field-of-view of Microsoft Azure Kinect in that distance can cover the body of all our subjects. Note that once the RGB-D camera can capture the whole body, the distance should be acceptable for our setup, but not strictly 1.5 meters. In addition, the number and placement of IMUs may also affect performance. Recently, some works demonstrate that with incomplete set or placement variation of IMUs, good kinematics estimation can still be achieved [64–66]. Hence, we believe that our system can perform well after some adaptations under these conditions. In the future, we will further evaluate the influence of RGB-D camera distance and different settings of IMUs on our system's performance.

Validation Scale. In this paper, we evaluate our system on our self-collected dataset with 19 subjects in total, including 10 athletes and 9 non-athlete subjects. Though the size of the dataset is comparable to other related research in biomechanics field [12, 18], the sample size of is still relatively small for justifying the system's deployability in a more real-world condition. However, as the person-wise performance of our system is relatively stable (see Fig. 10), we believe our system has a potential generalizability in a larger scale. Further exploration and validation will be left as a future work.

8 CONCLUSION

This paper presents ACLGuard, a physics-aware system for KAM estimation, designed to enhance ACL injury prevention during high-risk tasks. We observe that existing IMU-based solutions fall short due to insufficient body information and lack of ID principle encoding, and design a system integrating both motion and body information with respect to ID principle. Towards the challenge in obtaining body information and extracting related features, we propose a Biomechanical Information Retrieval module with a hybrid CNN-RNN module for kinematic feature extraction model and a MLP-based BSPs feature extraction model. For the challenge of KAM modeling based on ID principle with imperfect features, we present a Physics-aware KAM Modeling framework, encoding ID principle into attention-enhanced multi-task learning framework. Evaluation shows ACLGuard achieves an average RMSE of 0.176 Nm/kg and NRMSE of 11.5%, comparable to markerless motion capture systems while enabling on-field usability with a simpler setup. These results highlight ACLGuard's potential in real-world on-field training scenarios.

Acknowledgments

This research is supported in part by RGC under Contract CERG 16204523, 16205824, AoE/E-601/22-R, SRFS2425-6S05, R6021-20, and Contract R8015.

References

[1] 2024. Bertec Force Plates. https://www.bertec.com/products/force-plates.
[2] 2024. Noraxon 3D Motion Capture System. https://www.noraxon.com/our-products/3d-motion-capture-imu/.
[3] 2024. Vicon Motion Systems. https://www.vicon.com.
[4] Amelia JH Arundale, Holly J Silvers-Granelli, and Grethe Myklebust. 2022. ACL injury prevention: Where have we come from and where are we going? *Journal of Orthopaedic Research®* 40, 1 (2022), 43–54.
[5] Kevin Bill, Patrick Mai, Lasse Mausehund, Sigurd Solbakken, Tron Krosshaug, and Uwe G Kersting. 2024. Individualized Technique Feedback for Instant Technique Improvements and Knee Abduction Moment Reductions–A New Approach for 'Sidestepping'ACL Injuries? *International journal of sports physical therapy* 19, 5 (2024), 535.
[6] Scott Bonnette, Christopher A DiCesare, Adam W Kiefer, Michael A Riley, Kim D Barber Foss, Staci Thomas, Katie Kitchen, Jed A Diekfuss, and Gregory D Myer. 2019. Injury risk factors integrated into self-guided real-time biofeedback improves high-risk biomechanics. *Journal of sport rehabilitation* 28, 8 (2019), 831–839.
[7] Scott Bonnette, Christopher A DiCesare, Adam W Kiefer, Michael A Riley, Kim D Barber Foss, Staci Thomas, Katie Kitchen, Jed A Diekfuss, and Gregory D Myer. 2019. Real-time biofeedback is more effective than sham feedback for modifying high risk biomechanics. *Journal of sport rehabilitation* (2019), 1.
[8] Melissa A Boswell, Scott D Uhlrich, K Thomas, JA Kolesar, GE Gold, GS Beaupre, SL Delp, et al. 2021. A neural network to predict the knee adduction moment in patients with osteoarthritis using anatomical landmarks obtainable from 2D video analysis. *Osteoarthritis and cartilage* 29, 3 (2021), 346–356.
[9] Karen K Briggs, Jack Lysholm, Yelverton Tegner, William G Rodkey, Mininder S Kocher, and J Richard Steadman. 2009. The reliability, validity, and responsiveness of the Lysholm score and Tegner activity scale for anterior cruciate ligament injuries of the knee: 25 years later. *The American journal of sports medicine* 37, 5 (2009), 890–897.
[10] Shengze Cai, Zhiping Mao, Zhicheng Wang, Minglang Yin, and George Em Karniadakis. 2021. Physics-informed neural networks (PINNs) for fluid mechanics: A review. *Acta Mechanica Sinica* 37, 12 (2021), 1727–1738.
[11] Serena Cerfoglio, Manuela Galli, Marco Tarabini, Filippo Bertozzi, Chiarella Sforza, and Matteo Zago. 2021. Machine learning-based estimation of ground reaction forces and knee joint kinetics from inertial sensors while performing a vertical drop jump. *Sensors* 21, 22 (2021), 7709.
[12] Lionel Chia, Jordan T Andersen, Marnee J McKay, Justin Sullivan, Tomas Megalaa, and Evangelos Pappas. 2021. Evaluating the validity and reliability of inertial measurement units for determining knee and trunk kinematics during athletic landing and cutting movements. *Journal of Electromyography and Kinesiology* 60 (2021), 102589.
[13] Susmita Das, Amara Tariq, Thiago Santos, Sai Sandeep Kantareddy, and Imon Banerjee. 2023. Recurrent neural networks (RNNs): architectures, training tricks, and introduction to influential research. *Machine Learning for Brain Disorders* (2023), 117–138.
[14] Carlo J De Luca, L Donald Gilmore, Mikhail Kuznetsov, and Serge H Roy. 2010. Filtering the surface EMG signal: Movement artifact and baseline noise contamination. *Journal of biomechanics* 43, 8 (2010), 1573–1579.

[15] Scott L Delp, Frank C Anderson, Allison S Arnold, Peter Loan, Ayman Habib, Chand T John, Eran Guendelman, and Darryl G Thelen. 2007. OpenSim: open-source software to create and analyze dynamic simulations of movement. *IEEE transactions on biomedical engineering* 54, 11 (2007), 1940–1950.

[16] Stefano Di Paolo, Nicola Francesco Lopomo, Francesco Della Villa, Gabriele Paolini, Giulio Figari, Laura Bragonzoni, Alberto Grassi, and Stefano Zaffagnini. 2021. Rehabilitation and return to sport assessment after anterior cruciate ligament injury: quantifying joint kinematics during complex high-speed tasks through wearable sensors. *Sensors* 21, 7 (2021), 2331.

[17] Foivos I Diakogiannis, François Waldner, Peter Caccetta, and Chen Wu. 2020. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing* 162 (2020), 94–114.

[18] Bingfei Fan, Haisheng Xia, Junkai Xu, Qingguo Li, and Peter B Shull. 2021. IMU-based knee flexion, abduction and internal rotation estimation during drop landing and cutting tasks. *Journal of Biomechanics* 124 (2021), 110549.

[19] Kevin R Ford, Christopher A DiCesare, Gregory D Myer, and Timothy E Hewett. 2015. Real-time biofeedback to target risk of anterior cruciate ligament injury: a technical report for injury prevention and rehabilitation. *Journal of sport rehabilitation* 24, 2 (2015).

[20] Alli Gokeler, Romain Seil, Gino Kerkhoffs, and Evert Verhagen. 2018. A novel approach to enhance ACL injury prevention programs. *Journal of experimental orthopaedics* 5, 1 (2018), 22.

[21] Annamaria Guiotto, Alfredo Ciniglio, Fabiola Spolaor, Davide Pavan, Federica Cibin, Alex Scaldaferro, and Zimi Sawacha. 2021. Reliability and repeatability of ACL Quick Check®: a methodology for on field lower limb joint kinematics and kinetics assessment in sport applications. *Sensors* 22, 1 (2021), 259.

[22] Herbert Hatze. 2002. The fundamental problem of myoskeletal inverse dynamics and its implications. *Journal of biomechanics* 35, 1 (2002), 109–115.

[23] Becky Heinert, Drew Rutherford, and Thomas W Kernozek. 2022. Effectiveness of augmented feedback on drop landing using baseline vertical ground reaction in female athletes. *International Journal of Athletic Therapy and Training* 28, 1 (2022), 40–45.

[24] Timothy E Hewett, Gregory D Myer, Kevin R Ford, Mark V Paterno, and Carmen E Quatman. 2016. Mechanisms, prediction, and prevention of ACL injuries: Cut risk with three sharpened and validated tools. *Journal of Orthopaedic Research* 34, 11 (2016), 1843–1855.

[25] Md Sanzid Bin Hossain, Zhishan Guo, and Hwan Choi. 2023. Estimation of lower extremity joint moments and 3d ground reaction forces using imu sensors in multiple walking conditions: A deep learning approach. *IEEE Journal of Biomedical and Health Informatics* 27, 6 (2023), 2829–2840.

[26] Gerwyn Hughes. 2014. A review of recent perspectives on biomechanical risk factors associated with anterior cruciate ligament injury. *Research in sports medicine* 22, 2 (2014), 193–212.

[27] Yu Iwama, Kengo Harato, Shu Kobayashi, Yasuo Niki, Naomichi Ogihara, Morio Matsumoto, Masaya Nakamura, and Takeo Nagura. 2021. Estimation of the external knee adduction moment during gait using an inertial measurement unit in patients with knee osteoarthritis. *Sensors* 21, 4 (2021), 1418.

[28] Doyun Jung, Cheolwon Lee, and Heung Seok Jeon. 2024. Multi-Model Gait-Based KAM Prediction System Using LSTM-RNN and Wearable Devices. *Applied Sciences* 14, 22 (2024), 10721.

[29] Adam W Kiefer, Adam M Kushner, John Groene, Christopher Williams, Michael A Riley, and Gregory D Myer. 2015. A commentary on real-time biofeedback to augment neuromuscular training for ACL injury prevention in adolescent athletes. *Journal of sports science & medicine* 14, 1 (2015), 1.

[30] Pawel Kudzia, Erika Jackson, and Genevieve Dumas. 2022. Estimating body segment parameters from three-dimensional human body scans. *Plos one* 17, 1 (2022), e0262296.

[31] Myunghyun Lee and Sukyung Park. 2020. Estimation of three-dimensional lower limb kinetics data during walking using machine learning from a single IMU attached to the sacrum. *Sensors* 20, 21 (2020), 6277.

[32] Ilya Loshchilov, Frank Hutter, et al. 2017. Fixing weight decay regularization in adam. *arXiv preprint arXiv:1711.05101* 5 (2017).

[33] Urbano Lugrís, Manuel Pérez-Soto, Florian Michaud, and Javier Cuadrado. 2024. Human motion capture, reconstruction, and musculoskeletal analysis in real time. *Multibody System Dynamics* 60, 1 (2024), 3–25.

[34] Xiaolei Lv, Jinxiang Chai, and Shihong Xia. 2016. Data-driven inverse dynamics for human motion. *ACM Transactions on Graphics (TOG)* 35, 6 (2016), 1–12.

[35] Xiang Ma, Xuemei Li, Lexin Fang, Tianlong Zhao, and Caiming Zhang. 2024. U-mixer: An unet-mixer architecture with stationarity correction for time series forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 38. 14255–14262.

[36] Naureen Mahmood, Nima Ghorbani, Nikolaus F Troje, Gerard Pons-Moll, and Michael J Black. 2019. AMASS: Archive of motion capture as surface shapes. In *Proceedings of the IEEE/CVF international conference on computer vision*. 5442–5451.

[37] Jonas L Markström, Helena Grip, Lina Schelin, and Charlotte K Häger. 2019. Dynamic knee control and movement strategies in athletes and non-athletes in side hops: implications for knee injury. *Scandinavian journal of medicine & science in sports* 29, 8 (2019), 1181–1189.

[38] Microsoft. 2024. Azure Kinect DK. https://www.microsoft.com/en-us/d/azure-kinect-dk/8pp5vxmd9nhq?activetab=pivot:overviewtab.

[39] Dean D Molinaro, Inseung Kang, Jonathan Camargo, Matthew C Gombolay, and Aaron J Young. 2022. Subject-independent, biological hip moment estimation during multimodal overground ambulation using deep learning. *IEEE Transactions on Medical Robotics and Bionics* 4, 1 (2022), 219–229.

[40] Gregory D Myer, Kevin R Ford, Stephanie L Di Stasi, Kim D Barber Foss, Lyle J Micheli, and Timothy E Hewett. 2015. High knee abduction moments are common risk factors for patellofemoral pain (PFP) and anterior cruciate ligament (ACL) injury in girls: is PFP itself a predictor for subsequent ACL injury? *British journal of sports medicine* 49, 2 (2015), 118–122.

[41] Victoria Neilson, Sarah Ward, Patria Hume, Gwyn Lewis, and Andrew McDaid. 2019. Effects of augmented feedback on training jump landing tasks for ACL injury prevention: A systematic review and meta-analysis. *Physical Therapy in Sport* 39 (2019), 126–135.

[42] Agnethe Nilstad, Erich Petushek, Kam-Ming Mok, Roald Bahr, and Tron Krosshaug. 2023. Kiss goodbye to the 'kissing knees': No association between frontal plane inward knee motion and risk of future non-contact ACL injury in elite female athletes. *Sports biomechanics* 22, 1 (2023), 65–79.

[43] A Paszke. 2019. Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703* (2019).

[44] David J Pearsall, J Gavin Reid, and Lori A Livingston. 1996. Segmental inertial parameters of the human trunk as determined from computed tomography. *Annals of biomedical engineering* 24 (1996), 198–210.

[45] Jason S Pedley, Rhodri S Lloyd, Paul J Read, Isabel S Moore, Mark De Ste Croix, Gregory D Myer, and Jon L Oliver. 2020. Utility of kinetic and kinematic jumping and landing variables as predictors of injury risk: a systematic review. *Journal of Science in Sport and Exercise* 2 (2020), 287–304.

[46] Ethan Perez, Florian Strub, Harm De Vries, Vincent Dumoulin, and Aaron Courville. 2018. Film: Visual reasoning with a general conditioning layer. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 32.

[47] Apoorva Rajagopal, Christopher L Dembia, Matthew S DeMers, Denny D Delp, Jennifer L Hicks, and Scott L Delp. 2016. Full-body musculoskeletal model for muscle-driven simulation of human gait. *IEEE transactions on biomedical engineering* 63, 10 (2016), 2068–2079.

[48] Mohamed Irfan Mohamed Refai, Bert-Jan F Van Beijnum, Jaap H Buurke, and Peter H Veltink. 2020. Portable gait lab: estimating 3D GRF using a pelvis IMU in a foot IMU defined frame. *IEEE transactions on neural systems and rehabilitation engineering* 28, 6 (2020), 1308–1316.

[49] Per A Renström. 2013. Eight clinical conundrums relating to anterior cruciate ligament (ACL) injury in sport: recent evidence and a personal reflection. *British journal of sports medicine* 47, 6 (2013), 367–372.

[50] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 234–241.

[51] Patrick Sadoghi, Arvind von Keudell, and Patrick Vavken. 2012. Effectiveness of anterior cruciate ligament injury prevention training programs. *JBJS* 94, 9 (2012), 769–776.

[52] Anne Schmitz and Brian Noehren. 2014. What predicts the first peak of the knee adduction moment? *The Knee* 21, 6 (2014), 1077–1083.

[53] Haraldur B Sigurðsson, Jón Karlsson, Lynn Snyder-Mackler, and Kristín Briem. 2021. Kinematics observed during ACL injury are associated with large early peak knee abduction moments during a change of direction task in healthy adolescents. *Journal of Orthopaedic Research®* 39, 10 (2021), 2281–2290.

[54] Bernd J Stetter, Frieder C Krafft, Steffen Ringhof, Thorsten Stein, and Stefan Sell. 2020. A machine learning and wearable sensor based approach to estimate external knee flexion and adduction moments during various locomotion tasks. *Frontiers in bioengineering and biotechnology* 8 (2020), 9.

[55] Tian Tan, Anthony A Gatti, Bingfei Fan, Kevin G Shea, Seth L Sherman, Scott D Uhlrich, Jennifer L Hicks, Scott L Delp, Peter B Shull, and Akshay S Chaudhari. 2023. A scoping review of portable sensing for out-of-lab anterior cruciate ligament injury prevention and rehabilitation. *NPJ Digital Medicine* 6, 1 (2023), 46.

[56] Tian Tan, Peter B Shull, Jenifer L Hicks, Scott D Uhlrich, and Akshay S Chaudhari. 2024. Self-Supervised Learning Improves Accuracy and Data Efficiency for IMU-Based Ground Reaction Force Estimation. *IEEE Transactions on Biomedical Engineering* (2024).

[57] Tian Tan, Dianxin Wang, Peter B Shull, and Eni Halilaj. 2022. IMU and smartphone camera fusion for knee adduction and knee flexion moment estimation during walking. *IEEE Transactions on Industrial Informatics* 19, 2 (2022), 1445–1455.

[58] Tylan Templin, Christopher D Riehm, Travis Eliason, Tessa C Hulburt, Samuel T Kwak, Omar Medjaouri, David Chambers, Manish Anand, Kase Saylor, Gregory D Myer, et al. 2024. Evaluation of drop vertical jump kinematics and kinetics using 3D markerless motion capture in a large cohort. *Frontiers in Bioengineering and Biotechnology* 12 (2024), 1426677.

[59] Hsin-Ruey Tsai, Shih-Yao Wei, Jui-Chun Hsiao, Ting-Wei Chiu, Yi-Ping Lo, Chi-Feng Keng, Yi-Ping Hung, and Jin-Jong Chen. 2016. iKneeBraces: knee adduction moment evaluation measured by motion sensors in gait detection. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 386–391.

[60] Ryo Ueno, Alessandro Navacchia, Nathaniel A Bates, Nathan D Schilaty, Aaron J Krych, and Timothy E Hewett. 2020. Analysis of internal knee forces allows for the prediction of rupture events in a clinically relevant model of anterior cruciate ligament injuries. *Orthopaedic Journal of Sports Medicine* 8, 1 (2020), 2325967119893758.

[61] Chao Wang, Peter PK Chan, Ben MF Lam, Sizhong Wang, Janet H Zhang, Zoe YS Chan, Rosa HM Chan, Kevin KW Ho, and Roy TH Cheung. 2020. Real-time estimation of knee adduction moment for gait retraining in patients with knee osteoarthritis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28, 4 (2020), 888–894.

[62] Keenon Werling, Dalton Omens, Jeongseok Lee, Ioannis Exarchos, and C Karen Liu. 2021. Fast and feature-complete differentiable physics for articulated rigid bodies with contact. *arXiv preprint arXiv:2103.16021* (2021).

[63] Baichen Yang, Xinyi Zhang, Jiaxi Zhang, Zirui Huang, Qiqi Lu, Jin Zhang, Hai Hu, and Qian Zhang. 2024. KneeGuard: A Calibration-free Wearable Monitoring System for Knee Osteoarthritis Gait Re-training via Effortless Wearing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 4 (2024), 1–29.

[64] Xinyu Yi, Yuxiao Zhou, Marc Habermann, Soshi Shimada, Vladislav Golyanik, Christian Theobalt, and Feng Xu. 2022. 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*. 13167–13178.

[65] Xinyu Yi, Yuxiao Zhou, and Feng Xu. 2021. Transpose: Real-time 3d human translation and pose estimation with six inertial sensors. *ACM Transactions On Graphics (TOG)* 40, 4 (2021), 1–13.

[66] Chengxu Zuo, Yiming Wang, Lishuang Zhan, Shihui Guo, Xinyu Yi, Feng Xu, and Yipeng Qin. 2024. Loose inertial poser: Motion capture with IMU-attached loose-wear jacket. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2209–2219.