IMPROVING DATA EFFICIENCY AND ACCURACY OF IMU-DRIVEN BIOMECHANICAL ASSESSMENT VIA SELF-SUPERVISED LEARNING

Tian Tan[1*], Peter B. Shull[2], Akshay S. Chaudhari[1]
[1]Department of Radiology, Stanford University, Stanford, CA, USA
[2]School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, Shanghai, China
*Corresponding author's email: alanttan@stanford.edu

Introduction: Biomechanical measurement traditionally relies on force plates and marker-based motion tracking, confining the measurement environment and limiting the scale of biomechanical studies. Wearable inertial measurement units (IMUs) could enable human movement monitoring in natural environments in larger cohorts. Recently proposed deep learning techniques hold promise to enable IMU-driven gait assessment, however these techniques typically require a large amount of IMU data as well as synchronized force plate and marker data that serve as labels for model training. Such datasets are rare because their collection requires trained personnel to operate in-lab devices, thus costly and time-consuming. Self-supervised learning may mitigate this challenge of data scarcity. It aims to pre-train deep learning models for representation extraction based on a vast amount of "unlabeled" data. Then, the pre-trained models can be fine-tuned to various downstream estimation tasks with "labeled" data. In this study, we evaluated the benefits of self-supervised learning based on an "unlabeled" IMU dataset with 6.6 hours of human movements, which reduced requirements on the size of downstream "labeled" datasets by 60% - 90% for loading rate, ground reaction force, and knee flexion moment estimation.

Methods: An "unlabeled" dataset [1] consisting of 90 subjects performing 21 movements while wearing 17 IMUs was segmented into windows for self-supervised model pre-training. The accelerometer and gyroscope data from the same window contain mutual underlying information, and pre-training aims to learn to extract this relationship using a convolutional neural network (ResNet-50 [2]). To accomplish this, the accelerometer and gyroscope data were first fed into the model to generate their respective representations. The model was then optimized to increase the similarity between the accelerometer and gyroscope representations of the same window while minimizing those from different windows using a contrastive loss [3].

To evaluate the performance of the self-supervised model, we used three downstream datasets that are substantially different from pre-training dataset in terms of subject numbers, IMU numbers, and gold-standard measurements, i.e., using 1 IMU for estimating loading rate among 15 subjects [3], using 2 IMUs for estimating peak ground reaction force among 21 subjects [4], and using 8 IMUs for estimating peak knee flexion moment among 17 subjects [5]. The self-supervised model was fine-tuned on each downstream dataset to estimate the corresponding biomechanical parameter. It was also compared against a randomly-initialized model that is used in conventional supervised learning. To ensure a fair comparison, both models were evaluated via five-fold cross validation, with subjects randomly assigned to folds. For each fold, models were trained with the same number of sufficiently large training steps.

Results & Discussion: The self-supervised model matched the performance of a conventional randomly initialized model while reducing training data requirements by 90% for loading rate estimation, by 60% for peak ground reaction force estimation, and by 75% for peak knee flexion moment estimation (Fig. 1). When using the entire datasets [4] – [6] for training, the self-supervised model significantly improved the mean correlation coefficients from 0.87 to 0.91 for loading rate estimation, 0.79 to 0.83 for peak ground reaction force estimation, and 0.87 to 0.90 for peak knee flexion moment estimation (Fig. 1).

Although the pre-training dataset we used was larger than downstream datasets, scaling the size of pre-training dataset may further improve the effectiveness of self-supervised learning. Future research may consider using synthetic IMU data or simulated human movement to generate large-scale data for self-supervised learning.

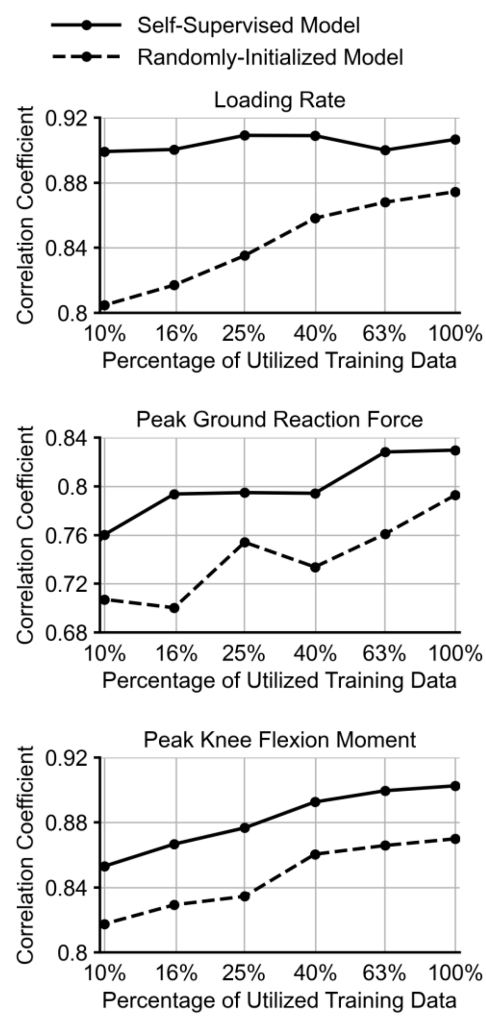

Figure 1: Correlation coefficients of biomechanical parameter estimation for a range of reduced training set sizes. Self-supervised models outperformed randomly-initialized models for all sizes.

Significance: Self-supervised learning with a large "unlabeled" IMU dataset for model pretraining can substantially improve data efficiency or improve the accuracy of deep learning. This approach could unlock newer and broader use cases of IMU-driven assessment where only limited "labeled" data is available.

Acknowledgements: This work was supported by the Joe and Clara Tsai Foundation through the Wu Tsai Human Performance Alliance.

References: [1] Ghorbani et al. (2021), *Plos One* 16(6); [2] He et al. (2016) *IEEE CVPR*; [3] Alayrac et al. (2020) *NIPS*; [4] Tan et al. (2021), *IEEE JBHI* 25(4); [5] Camargo et al. (2021), *J Biomech* 119; [6] Tan et al. (2022), *IEEE TII* 19(2).