Foot impact force estimation from monocular video during gait using computer vision and machine learning

Pawel Kudzia
School of Biomedical Engineering
Email: pawel.kudzia@ubc.ca

Summary
Measuring ground reaction forces (GRFs) is important for injury prevention and rehabilitation. We developed a video-based method to estimate leg-specific vertical GRFs during gait, reducing the need for lab equipment. Thirty participants walked at three speeds while we recorded synchronized GRF and video data. Using a computer vision approach, we estimated sparse kinematic landmarks and trained a transformer-based neural network to predict vertical GRFs. Our preliminary model evaluation showed high accuracy ($R^2 = 0.94 \pm 0.06$, RMSE = 0.06 ± 0.03 BW). Future work will refine this approach for portable biomechanical assessments.

Introduction
Accurately measuring ground reaction forces is essential for assessing gait mechanics, diagnosing disorders, and improving rehabilitation strategies. Researchers increasingly use video-based analysis to evaluate movement, reducing the need for laboratory equipment. Computer vision advances enable researchers to extract movement data from video efficiently and at a low computational cost. Preliminary studies have successfully estimated forces directly from video-captured motion, yielding promising results [1]. In this study, we aim to achieve two main objectives: first, to develop a system that estimates GRFs independently for each leg during walking at different speeds; and second, to analyze gait dynamics using these GRF estimates.

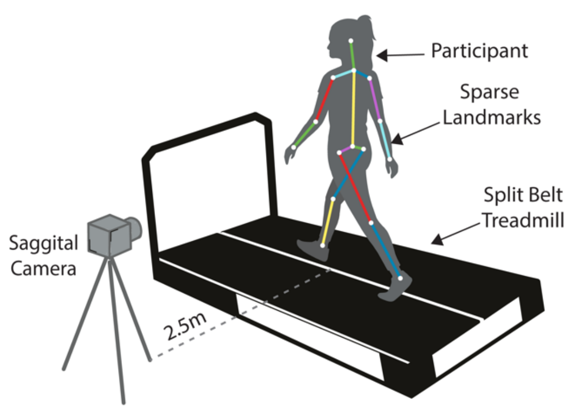

Figure 1: Experimental setup.

Methods
We recruited 30 participants (18F, 11M) and instructed them to walk at three speeds (0.8, 1.2, and 1.6 m/s). During each walking trial, participants walked for 2 minutes per speed while we recorded synchronized GRF measurements (800 Hz) and RGB video (80 Hz) (Fig 1). We extracted kinematic data from each video using OpenPose, a computer vision algorithm that identifies 25 sparse kinematic landmarks on the body in each video frame [2]. Using these kinematic data, we constructed a dataset to train our model. We developed a neural network with a transformer-based architecture, where the model used sparse kinematic data as input to predict vertical GRFs for each leg as output. To evaluate accuracy and generalizability, we applied the leave-one-participant-out cross-validation method, training the model on data from all but one participant and testing it on the excluded participant, repeating this process for all participants. We assessed model performance by comparing its predictions to empirical data using R^2 and RMSE.

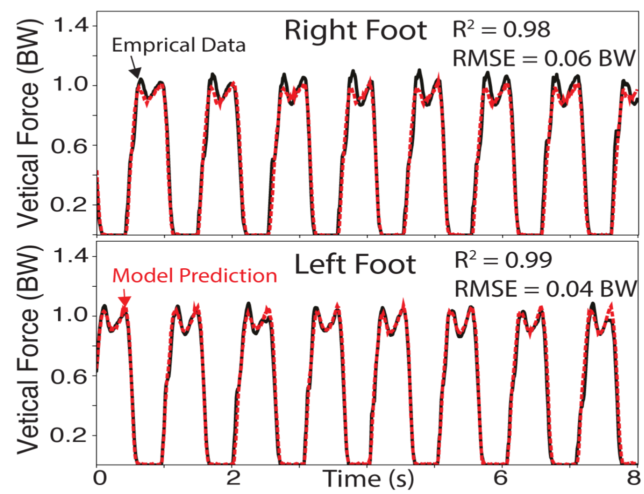

Figure 2: Example GRF prediction for right and left foot.

Results and Discussion Our preliminary results indicate that our model predicts vertical GRFs with an R^2 of 0.94 ± 0.06 and an RMSE of 0.06 ± 0.03 body weight (BW) on average for the right leg, with similar accuracy for the left leg (Fig 2). Our next step is to test different model architectures and determine the accuracy of predicted gait dynamics.

Conclusions
We developed a video-based method to estimate vertical GRFs for each leg, reducing the need for specialized equipment. Our long-term vision is to refine this technique to create a portable system for measuring biomechanical metrics traditionally confined to the lab.

Acknowledgments

This data was collected at the SFU Locomotion Lab.

References

[1] T. B. Aderinola et al. IEEE Open J Eng Med Biol, vol. 4, pp. 109–115, 2023.
[2] Z. Cao et al. IEEE Trans. Pattern Anal. Mach. Intell. 2019.
[3] M. Mundt et al. Sensors , vol. 23, no. 1, Dec. 2022.