PREDICTING GROUND REACTION FORCES FROM 2D VIDEO: BRIDGING THE LAB TO FIELD NEXUS

Molly Goldacre[1,4], Amar El-Sallam[2], Hannah Wyatt[3], William Johnson[5], Jian Liu[2], Ajmal Mian[2] and Jacqueline Alderson[1,3]

Minderoo Tech & Policy Lab, The University of Western Australia, Perth, Australia[1]
School of School of Physics, Mathematics & Computing, University of Western Australia, Perth, Australia[2]
Auckland University of Technology, Sports Performance Research Institute New Zealand (SPRINZ), Auckland, New Zealand[3]
School of Human Sciences, University of Western Australia, Perth, Australia[4]
Houston Astros Baseball Club, Texas, USA[5]

The accurate measurement of ground reaction forces (GRFs) is confined to laboratory-based settings, posing an ongoing problem for sports biomechanists working in the field given the necessity of this variable for a variety of biomechanical modelling approaches. The ability to remote estimate on-field GRFs could facilitate the development of a tool that could positively impact the side-line management of athletes during match play. Using laboratory collected GRF side stepping and running data, alongside concurrently collected two-dimensional (2D) video, the aim of this study was to use a least squares estimator (LSE) matrix to estimate GRFs from 2D video. Results of r>0.8 were found for the vertical and horizontal GRF components which was slightly lower than the r>0.9 observed for the higher complexity convolutional neural network (CNN) approach which was used as a comparator model. These results provide early support for the efficacy of remote on-field estimation of GRFs determined from 2D video footage in isolation.

KEY WORDS: prediction, ground reaction forces, least squares estimator, non-invasive

INTRODUCTION: Anterior cruciate ligament (ACL) rupture is one of the most debilitating injuries that an athlete of any playing level, can sustain and is most commonly a career-ending event (Dallalana et al., 2007). The ACL injury mechanism can be considered an excessive load related event (Lloyd & Buchanan 2001). Alongside movement technique factors, studies have shown high knee joint moments, specifically external knee valgus (abduction) and internal rotation moments during unplanned sidestepping tasks, will increase ACL injury risk (Dallalana et al., 2007). Traditional estimation of joint loads during unplanned sidestepping tasks relies on an inverse dynamics approach that requires the direct recording of the forces between the foot and the ground. To achieve this, researchers can either bring athletes into a laboratory environment where ground embedded force plates directly record all components of the GRFs and moments (Besier et al., 2001), or instrument an athletes' shoes with force transducers, or other sensors, to directly record or estimate these forces (Debbi et al. 2012, Price et al., 2016). Both approaches limit the ecological validity of collecting such data as they result in interference to the environment or successful task completion. Recent advances in the estimation of GRFs from wearable sensors (Karatsidis et al., 2016) are promising, however, these have only been tested in low velocity walking gait scenarios, have only estimated the vertical force component (F_z) (Wundersitz et al., 2013), and are limited in application by low-fidelity, low-resolution capture whilst overfitting simple movement patterns (Camomilla et al., 2018, Johnson et al, 2018 a,b). Recent advances in computer vision and data analytics provides an opportunity to bypass the traditional GRF data collection limitations. Johnson and colleagues (2018 a,b) recently used a pre-trained CaffeNet convolutional neural network (CNN) to estimate three-dimensional (3D) GRFs and moments for sidestepping. This work, while one of the first papers to establish the potential for machine learning to predict biomechanical data from related variables, is still reliant on laboratory collected marker data. The aim of the present study is to create a novel LSE matrix approach to predict GRFs from standard commercial grade 2D video, and to evaluate the lower computational method to a higher complexity CNN method.

METHODS: This study comprised of an experimental data collection phase and a computer model development phase. The study design was such that the LSE estimated GRFs were compared to the ground truth force plate data, and to assess model performance, was also compared to the CNN model of Johnson et al., (2018) which required traditional three-dimensional (3D) trajectory marker inputs.

Fifteen semi-professional and amateur female Australian Rules Football players (23 ± 3.7 years, 62.7 ± 5.4 kilograms) attended a testing session at the University of Western Australia's (UWA) sports biomechanics laboratory. All participants were injury free and provided informed consent prior to testing in accordance with UWA Human Ethics approval (RA/4/1/2593). A 23 camera Vicon MX system (Vicon Peak, Oxford Metrics, Oxford, UK) collecting at 200Hz and synchronised with a 1.2 m x 1.2 m AMTI force plate (Advanced Mechanical Technology Inc., Watertown, MA) sampling at 2000 Hz, was used to capture 3D motion and force data. Three high-definition 2D video cameras (Sony HDR-CX700, 50 Hz) captured all trials and were positioned directly to the right of the force platform in the following locations: a) a true sagittal (TS) view, b) slightly anterior to true sagittal – named anterior sagittal (AS) view, and c) slightly posterior sagittal (PS) to the true sagittal view. Participants were affixed with 67 retro-reflective markers as per UWA's custom marker set and model, however only eight of these markers were necessary inputs required by the CNN model of Johnson et al. (2018) (Caffenet). Participants performed a randomised series of sidesteps, crossover sidesteps and straight-line runs. Timing gates were positioned three metres and five metres from the posterior edge of the force plate to trigger a large arrow task type stimulus when the participants were 0.5 m from the force plate. A total of 476 successful 3D motion capture trials were captured and 90% (n=428) were fed into the pre-trained CaffeNet prediction model of Johnson et al. (2018). Once passed through, the correlation (*r*) values between the predicted and ground truth waveforms were used as the comparison output to the new LSE method.

An overview of the computer model development phase is presented in Figure 1. To begin 1,428 2D videos passed through data hygiene checks and were labelled according to camera view, trial number and participant. For each video, frames were extracted as individual *.png* images using an FFmpeg script in MATLAB. Manual visual detection of foot strike (FS) and the foot off (FO) event frames was completed for all videos to identify stance phase, a process serving to sync and temporally normalise the video and force plate data data. Prior to developing the LSE matrix a *surrogate* projected 2D centre of mass (COM) was identified in each .png frame using a weighted image centroid calculation. This was to establish the correlation correspondence of the vertical and horizontal components (C_x, C_y) of the centroid and the $GRF_{x,y,z}$ waveforms. 2D surrogate COM trajectories for each individual frame were segmented from the background using an average frame padding technique in MATLAB.

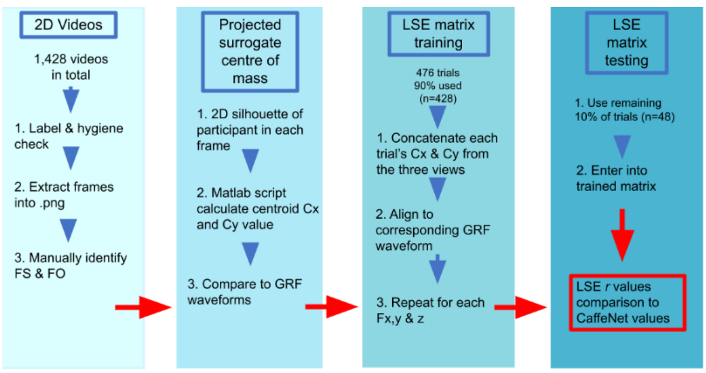

Figure 1. Computer model development: 2D video processing, through surrogate COM projection and the training and testing of the LSE matrix.

Custom MATLAB script calculated the average weighted image centroid and output C_x and C_y coordinate values for each silhouette frame. The LSE matrix was trained using 90% of the available 3D marker trajectory trials (n=428) by concatenating each trial's AS, TS and PS camera view 2D surrogate COM (C_x and C_y) coordinate waveforms to the corresponding $GRF_{x,y,z}$ waveforms. This assembled the computational matrix for the LSE using non-linear regression. Transformation required the centroid and GRF matrices to be unit vector normalised (same dimension, length and height). GRF data vectors were down sampled by a factor of 10 to match the sample length of the centroid values. Centroid matrices were then repeated for each F_x , F_y, and F_z components, with separate and unique matrices computed for each waveform. The training set was passed through the LSE model to extract the vector features to estimate F_x , F_y, and F_z independently. The final 10% (n=48) of the concatenated trial centroids remained unseen until passed through the LSE model, thereby acting as the test validation set. Predicted waveforms of the test set were then compared with the Pearson correlation coefficients (r) calculated for each sidestepping and running trial estimated and ground truth $GRF_{x,y,z}$(mean). LSE performance (r) was also compared against those obtained using the pretrained CaffeNet model from Johnson et al., (2018).

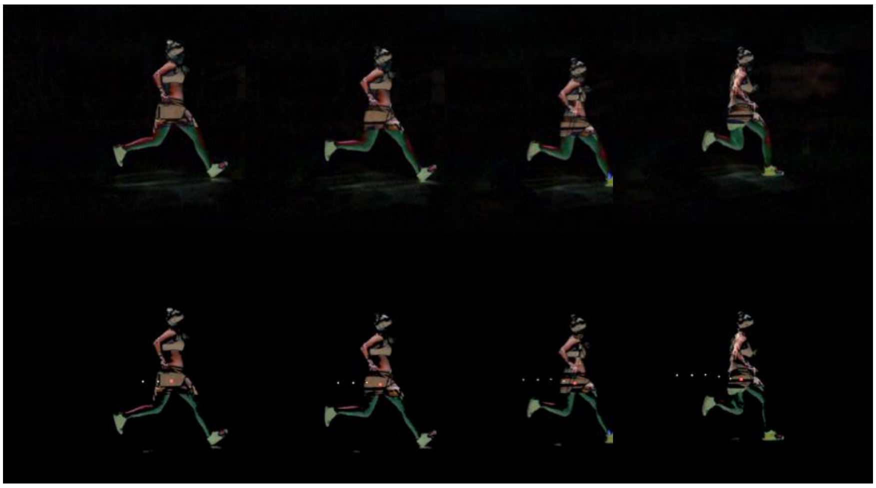

Figure 2. Silhouette Segmentation – Original segmented frames (top), segmented frames (bottom) with red dot denoting the estimated 2D surrogate COM and the small white dots denoting the tracking of its C_x, C_y trajectory components throughout consecutive representative frames of a trial.

RESULT: The deep CNN pre-trained CaffeNet model of Johnson et al., (2018) reported strong correlation coefficients ranging from (0.75-0.946) for prediction of all $GRF_{x, y, z}$ components (Table 1). Predicted $GRF_{x, y, z}$ components from application of the LSE model reported strong positive *r* values ranging from 0.806-0.903 for the F_y and F_z forces. LSE model predictions for F_x derived from running and sidestepping concatenated centroids were poor (0.263 running; -0.114 sidestepping) which was not unexpected given stage one results which failed to establish a relationship between the $C_{x,y}$ trajectories and the F_x force component waveform. The LSE method was implemented to find a computational matrix which was able to predict F_y and F_z waveforms with correlation coefficients above 0.800 for both running and sidestepping manoeuvres.

DISCUSSION: The LSE estimated values were below the CaffeNet model results (coefficients above 0.900) which for the purpose of this study served as the GRF prediction model comparison. By comparison, the LSE method developed in this study was a much-simplified machine learning matrix approach. The results show that a lower computational complexity method is able to perform to a similar standard as a deep learning model, likely due to the nature of the feature vectors that were entered into the LSE matrix to be learned by the model (i.e. surrogate COM centroids). A benefit of using the LSE method was that control over which exact features were passed to the machine learning model were the sole input features (i.e. known meaningful COM surrogate coordinates). The poorer results of the LSE prediction for medio-lateral forces compared with the CaffeNet model suggests that the video derived 2D surrogate COM as a single feature vector for training a machine learning model is not adequate for the prediction of F_x (medio-lateral) forces.

	Ground Reaction Forces		
	Fy (*r*)	Fz (*r*)	Fx (*r*)
LSE			
Sidestepping	0.903 (± 0.076)	0.865 (± 0.217)	-0.114 (± 0.775)
Running	0.806 (± 0.135)	0.816 (± 0.251	0.263 (± 0.829)
CaffeNet (Johnson et al., 2012)			
Sidestepping	0.946	0.942	0.930
Running	0.930	0.750	0.484

Table 1. LSE and CaffeNet mean correlation coefficient (*r*) for F_x, F_y and F_z predictions compared with ground truth data. Sidestepping trial n=158; running trial n= 318 for both model analyses.

CONCLUSION: The LSE method was implemented to create feature vectors and a computational matrix to predict anterior-posterior and medial-lateral ground reaction forces. Prediction results were slightly poorer using a simple LSE approach when compared with the more complex CNN CaffeNet model, however this study provides initial evidence for the ability of lower complexity methods to produce similar strength results without the need for additional processing power or computational time. With further fine-tuning, feature engineering and increased trial input data these results will likely improve. The overall significance and practical relevance of these results establishes early-stage efficacy of markerless machine learning methods to predict GRF components from 2D video. This exploratory investigation suggests that methods with lower computational complexity, such as the LSE model, are a viable option for non-invasive, 2D video-based prediction of key GRF components.

REFERENCES

Besier T., Lloyd D., Cochrane J., Ackland T. (2001). External loading of the knee joint during running and cutting maneuvers. *Medicine & Science in Sports & Exercise* 33(7), 1168-1175.

Camomilla V., Bergamini E., Fantozzi S., Vannozzi G. (2018). Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: A systematic review. *Sensors* 18(3), 873.

Dallalana R., Brooks J., Kemp S., Williams A. (2007). The epidemiology of knee injuries in English professional rugby union. *The American Journal of Sports Medicine* 35(5), 818-830.

Debbi E., Wolf A., Goryachev Y., Yizhar Z., Luger E., Debi R., Haim A. (2012). In-shoe center of pressure: Indirect force plate vs. direct insole measurement. *The Foot* 22(4), 269-275.

Johnson W., Alderson J., Lloyd D., Mian A. (2018a). Predicting Athlete Ground Reaction Forces and Moments from Spatio-temporal Driven CNN Models. *IEEE Transactions on Biomedical Engineering* 66(3), 689-694.

Johnson W., Mian A., Donnelly C., Lloyd D., Alderson J. (2018b). Predicting athlete ground reaction forces and moments from motion capture. *Medical and Biological Engineering and Computing* 56(10), 1781-1792.

Karatsidis A., Bellusci G., Schepers H., de Zee M., Andersen M., Veltink P. (2016). Estimation of ground reaction forces and moments during gait using only inertial motion capture. *Sensors* 17(1), 75.

Lloyd D., Buchanan T. (2001). Strategies of muscular support of varus and valgus isometric loads at the human knee. *Journal of Biomechanics* 34(10), 1257-1267.

Price C., Parker D., Nester C. (2016). Validity and repeatability of three in-shoe pressure measurement systems. *Gait & Posture* 46:69-74.

Wundersitz D., Netto K., Aisbett B., Gastin P. (2013). Validity of an upper-body-mounted accelerometer to measure peak vertical and resultant force during running and change-of-direction tasks. *Sports Biomechanics* 12(4), 403-412.