ESTIMATING THE PEAK VERTICAL GROUND REACTION FORCE COMPONENT AND STEP TIME IN TREADMILL RUNNING USING MACHINE LEARNING - A PILOT STUDY

Alex Ong[1], Ian Harris Sujae[1] and Joseph Hamill[2]

School of Sports, Health and Leisure, Republic Polytechnic, Singapore[1]
School of Public Health & Health Sciences, University of Massachusetts Amherst, USA[2]

This study aims to investigate the efficacy of a stacking approach to estimate parameters in treadmill running. Nineteen participants ran on a treadmill at self-selected pace. Ground reaction force and kinematic data were collected. Stacking in machine learning was used to estimate the peak vertical ground reaction force and step time. Good agreement was observed in the test data set for predicted and measured values of the peak vertical ground reaction force component and step time where the ICC values were 0.85 and 0.99 respectively. This suggests stacking may be a feasible approach to estimate kinetic and kinematic parameters during treadmill running.

KEYWORDS: stacking, data science, machine learning, vertical ground reaction force, step time, treadmill running

INTRODUCTION: The measurement of vertical ground reaction forces ($vGRF_{peak}$) and spatial parameters complement treadmill running analysis (Arnold et al., 2019). Such a setup for measuring $vGRF_{peak}$ directly using force-transducers instrumentation is costly and potentially limits its adoption (Hong et al., 2017). There are studies that indirectly estimate ground reaction forces (GRF) during gait without the use of force transducers (Onal et al., 2019). Additionally, spatial parameters such as step and stride time have been estimated (Zhang et al., 2019). The use of machine learning in data science for such estimations is an emerging area of investigation (Burdack et al., 2019). However, these types of investigations typically require 'large data' to optimise ecological validity (Ferber et al., 2016). Recently, machine learning approaches have been applied to small data set that require models with low complexity to avoid overfitting the model to the data (Lim et al., 2020).

Stacking in ensemble-based machine learning is one approach that addresses overfitting (Alizadeh et al., 2019) as it has shown potential benefits in obtaining accurate results (Mekruksavanich et al., 2019) as it leverages on the strengths of various models . In treadmill activities, the motion is constrained within a small spatial volume and the activities are repetitive (Ong et al., 2017). This is potentially advantageous in a simple model as compared to a complex movement model (e.g., dancing and cutting maneuvers). However, there is a dearth of knowledge on estimating kinetic and spatial parameters using machine learning stacking approach.

Therefore, this pilot study aims to investigate the efficacy of stacking approach to estimate $vGRF_{peak}$ and step time in treadmill running. It is hypothesized that the predicted $vGRF_{peak}$ and step time have good ICC measures between the actual measured and predicted values.

METHODS: Nineteen healthy female participants (age = 33.78 ± 6.20 yr, height = 1.60 ± 0.06 m, mass = 55.56 ± 7.56 kg) were included in the study. The study was approved by the institutional review board and written consent was obtained from the participants prior to the study. Four Codamotion sensors (model: Coda CX1) were positioned around a Gaitway instrumented treadmill (model: H/P Cosmos Gaitway II S). The Codamotion system (data captured at 200Hz) was synchronized with the Gaitway instrumented treadmill (data captured at 1KHz). In this investigation, the participants performed jogging trials, in their own shoes, on the instrumented treadmill at self-selected speeds at a 0% gradient. Active markers were placed on the iliac crests as well as on the lateral aspects of femur greater trochanter, femur medial epicondyle, fibula apex of medial malleolus, head of 5th metatarsal of both the left and right lower limbs. The participant jogged 6 minutes for familiarization on the treadmill (Meyeret al., 2019) and was allowed to make finer adjustments of speed during the jog. After the 6 minutes, the kinematic data of the pelvis, lower limbs as well as the $vGRF_{peak}$ (normalized to body weight, BW) and step times values were collected for 10 strides (2 steps per stride). The time-series kinematic data set (comprising of the markers' 3D-coordinates from 19 participants x 20 steps; 380 samples) was divided with a 60-20-20 split (sample-wise) into training, validation, and test data. Python 3.7 via the Orange data mining toolkit was used to facilitate the analysis (Demšar et al., 2013) using its default configuration tuning values. The data were trained and evaluated using stacking where the base predictors were random forest, linear regression, neural network and the meta learner was a neural network (Figure 1). Intra-class correlation (ICC; *Two-way mixed average measures/ absolute agreement* model) between the predicted and actual measured data as well as the root mean square errors (RMSE) and the Bland and Altman plots (Bland & Altman, 1986) with limits of agreement were used to investigate the efficacy of the approach.

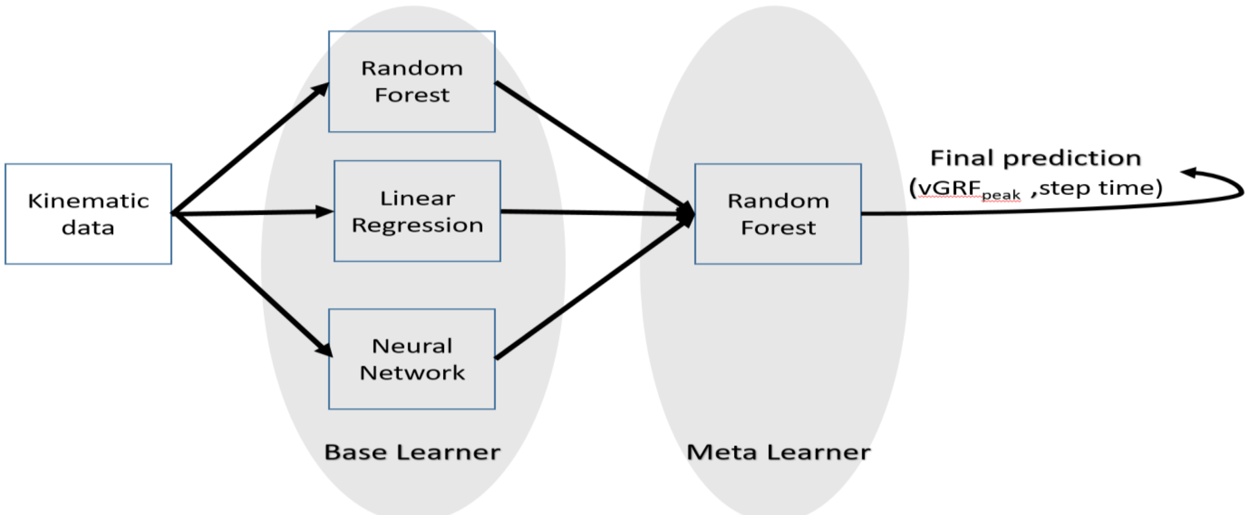

Figure 1: A schematic diagram of the stacking machine learning model.

RESULTS: The self-selected jogging rate was 7.2 ± 0.7 km/h. There were no significant differences found in both the left and right feet for both the kinetic and kinematic data ($p>0.05$). Thus, the data of both the left and right limbs were pooled in the analysis. The ICC measures were good for both $vGRF_{peak}$ (>0.85) and step times (>0.98) (Table 1). In the Bland and Altman plot (Figure 2) with mean (bias) of 0.23 BW, the linear regression line quantifying the differences between the measured and predicted $vGRF_{peak}$ with a slope of 0.04 ($p=0.64$).

Table 1: Mean, standard deviation (SD), root mean square error (RMSE) and intra-class correlation (ICC) values of results from training, validation and test data. The results were computed from the measured $vGRF_{peak}$ and Step time values (*Reference*) with the predicted values from random forest (RF), linear regression (LR) and neural network (NN) algorithms for Base Learner and using RF for Meta Learner.

		Training				Validation				Test	
		Reference	*RF*	*LR*	*NN*	*Reference*	*RF*	*LR*	*NN*	*Reference*	*RF*
$vGRF_{peak}$	Mean /BW (SD)	1.645 (0.219)	1.615 (0.219)	1.644 (0.328)	1.618 (0.241)	1.635 (0.200)	1.595 (0.219)	1.551 (0.299)	1.571 (0.241)	1.651 (0.232)	1.511 (0.300)
	RMSE /BW	-	0.155	0.423	0.213	-	0.215	0.253	0.200	-	0.242
	ICC	-	0.999	0.802	0.927	-	0.892	0.713	0.898	-	0.852
Step Time	Mean / s (SD)	0.382 (0.081)	0.378 (0.020)	0.360 (0.011)	0.372 (0.018)	0.373 (0.072)	0.370 (0.012)	0.369 (0.021)	0.368 (0.021)	0.381 (0.068)	0.379 (0.020)
	RMSE/ s	-	0.028	0.024	0.030	-	0.022	0.038	0.021	-	0.021
	ICC	-	0.999	0.999	0.999	-	0.989	0.991	0.989	-	0.989

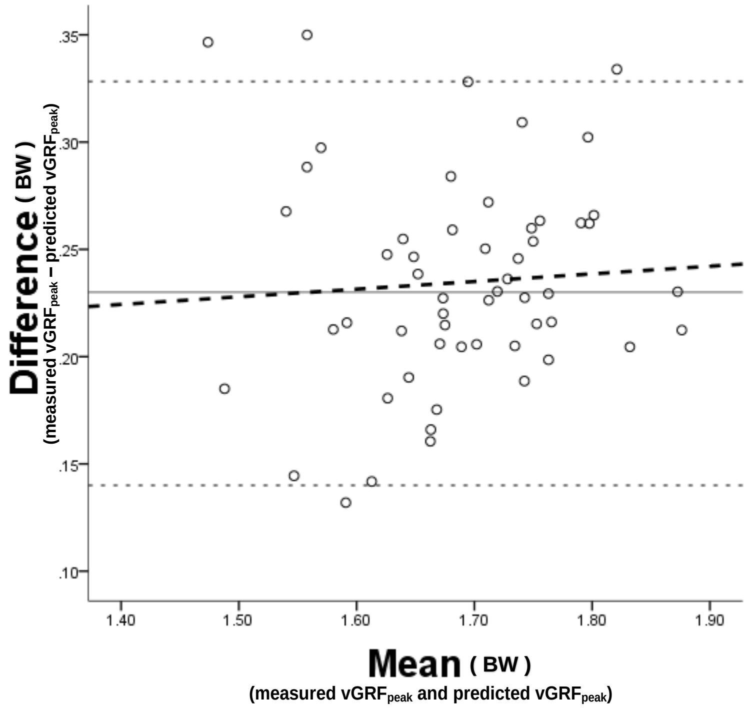

Figure 2: Bland and Altman plot of the differences and means of measured $vGRF_{peak}$ and predicted $vGRF_{peak}$ normalised to body weight (BW). The mean and the limits (computed at 1.96 X standard deviation) are depicted with solid and dotted horizontal lines respectively. The regression line quantifying the differences between the measured and predicted $vGRF_{peak}$ is included as a dashed line.

DISCUSSION: This study aimed to investigate the efficacy of stacking approach to estimate $vGRF_{peak}$ and step time in treadmill running. The main finding was that good ICC measures of agreement were observed between predicted and measured values for $vGRF_{peak}$ and step times. Therefore, we accepted our hypothesis. The step times (spatial parameters) have good ICC measures as well as low RMSE values (Table 1). This is potentially useful for estimating cadence during treadmill analyses where cadence is 1/(2 x step time). From here, we may estimate the stride length where stride length is speed/(cadence) and which the gait speed is typically known during treadmill analyses. The observed good ICC values are consistent with other existing methods such as using inertia sensors (Washabaugh et al., 2017).
The $vGRF_{peak}$ estimates have good ICC measures with an average of 0.2 BW underestimation of the actual reference measured values (Figure 1). This approach potentially allows a simpler method to model the machine learning algorithm in this study where the RMSE was 0.21 BW or approximately 13% of mean $vGRF_{peak}$. Recent studies prediction GRF have a mean 14% percent error over a range of gait speeds (Nedergaard et al., 2018; Miller et al., 2019). RMSE may improve with a more complex model approach but this approach potentially requires larger data sets and more computational resources.
The machine learning based stacking approach was tuned using the default configurations of the toolkit. Other tuning methods exist such as automatic and random but they are computationally costly (Happ et al., 2012). However, with technological advances deriving faster processing speeds such approaches will be potentially more feasible (Silla et al., 2020).

CONCLUSION: The results suggest that machine learning approaches in data science may be used to estimate kinetic ($vGRF_{peak}$) and spatial (step time) parameters. This may be a feasible alternative in treadmill gait analysis in clinical settings. The observations of estimated kinetic

($vGRF_{peak}$) and spatial (step time) parameters could provide meaningful feedback for the clinician to develop rehabilitation protocols. We are not suggesting that such an approach replaces the sophisticated instrumented treadmill systems. When more detailed quantitative measurements are required, treadmills with force transducer systems should be used.
For this pilot study, to investigate the efficacy of stacking approach in data science to estimate gait parameters during treadmill gait analysis, we used a relatively small data set. Additionally, this study involved only females at self-selected pace. Thus, future studies with more participants of both genders as well as varying gait speeds and gradients can be conducted to investigate the ecological validity of this stacking approach.

REFERENCES

Alizadeh, R., Jia, L., Nellippallil, A. B., Wang, G., Hao, J., Allen, J. K., & Mistree, F. (2019). Ensemble of surrogates and cross-validation for rapid and accurate predictions using small data sets. *AI EDAM*, 1-18.

Arnold, B. J., Weeks, B. K., & Horan, S. A. (2019). An examination of treadmill running familiarisation in barefoot and shod conditions in healthy men. *Journal of sports sciences*, *37*(1), 5-12.

Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, *327*(8476), 307-310.

Burdack, J., Horst, F., Giesselbach, S., Hassan, I., Daffner, S., & Schöllhorn, W. I. (2019). Systematic Comparison of the Influence of Different Data Preprocessing Methods on the Classification of Gait Using Machine Learning. *arXiv preprint arXiv:1911.04335*.

Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevar, T., Milutinovič, M., & Štajdohar, M. (2013). Orange: data mining toolbox in Python. *The Journal of Machine Learning Research*, *14*(1), 2349-2353.

Ferber, R., Osis, S. T., Hicks, J. L., & Delp, S. L. (2016). Gait biomechanics in the era of data science. *Journal of biomechanics*, *49*(16), 3759-3761.

Happ, P., Feitosa, R. Q., & Street, A. (2012). Assessment of optimization methods for automatic tuning of segmentation parameters. *Proceedings of Geographic object-based image analysis (GEOBIA).*

Hong, C. Y., Guo, L. Y., Song, R., Nagurka, M. L., Sung, J. L., & Yen, C. W. (2017). Developing a Low-Cost Force Treadmill via Dynamic Modeling. *Journal of Healthcare Engineering*, *2017*.

Lim, H., Kim, B., & Park, S. (2020). Prediction of Lower Limb Kinetics and Kinematics during Walking by a Single IMU on the Lower Back Using Machine Learning. *Sensors*, *20*(1), 130.

Mekruksavanich, S., & Jitpattanakul, A. Classification of Gait Pattern with Wearable Sensing Data. (2019). *Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT-NCON),* pp. 137-141, IEEE.

Meyer, C., Killeen, T., Easthope, C. S., Curt, A., Bolliger, M., Linnebank, M., & Filli, L. (2019). Familiarization with treadmill walking: How much is enough?. *Scientific reports*, *9*(1), 5232.

Miller, A. C., Aljohani, M., Kim, H., & Kipp, K. (2019). A neural network method to Predicting stance phase ground reaction force in distance runners. *ISBS Proceedings Archive*, *37*(1), 399.

Nedergaard, N. J., Verheul, J., Drust, B., Etchells, T., Lisboa, P., Robinson, M. A., & Vanrenterghem, J. (2018). The feasibility of predicting ground reaction forces during running from a trunk accelerometry driven mass-spring-damper model. *Peer J*, *6*, e6105.

Onal, S., Leefers, M., Smith, B., & Cho, S. (2019). Predicting running injury using kinematic and kinetic parameters generated by an optical motion capture system. *SN Applied Sciences*, *1*(7), 675.

Ong, A., Harris, I. S., & Hamill, J. (2017). The efficacy of a video-based marker-less tracking system for gait analysis. *Computer methods in biomechanics and biomedical engineering*, *20*(10), 1089-1095.

Silla, F., Prades, J., Baydal, E., & Reaño, C. (2020). Improving the performance of physics applications in atom-based clusters with rCUDA. *Journal of Parallel and Distributed Computing*, *137*, 160-178.

Washabaugh, E. P., Kalyanaraman, T., Adamczyk, P. G., Claflin, E. S., & Krishnan, C. (2017). Validity and repeatability of inertial measurement units for measuring gait parameters. *Gait & posture*, *55*, 87-93.

Zhang, H., Guo, Y., & Zanotto, D. (2019). Accurate Ambulatory Gait Analysis in Walking and Running Using Machine Learning Models. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*.

ACKNOWLEDGEMENTS: The authors would like to thank all the research assistants and laboratory technicians who have contributed to the successful completion of this study.