

Individualized learning-based ground reaction force estimation in people post-stroke using pressure insoles

Gregoire Bergamo^{1*}, Krithika Swaminathan^{1*}, Daekyum Kim^{1*}, Andrew Chin¹, Christopher Sivi¹, Ignacio Novillo¹, Teresa C. Baker², Nicholas Wendel², Terry D. Ellis², and Conor J. Walsh^{1†}

Abstract—Stroke is a leading cause of gait disability that leads to a loss of independence and overall quality of life. The field of clinical biomechanics aims to study how best to provide rehabilitation given an individual’s impairments. However, there remains a disconnect between assessment tools used in biomechanical analysis and in clinics. In particular, 3-dimensional ground reaction forces (3D GRFs) are used to quantify key gait characteristics, but require lab-based equipment, such as force plates. Recent efforts have shown that wearable sensors, such as pressure insoles, can estimate GRFs in real-world environments. However, there is limited understanding of how these methods perform in people post-stroke, where gait is highly heterogeneous. Here, we evaluate three subject-specific machine learning approaches to estimate 3D GRFs with pressure insoles in people post-stroke across varying speeds. We find that a Convolutional Neural Network-based approach achieves the lowest estimation errors of 0.75 ± 0.24 , 1.13 ± 0.54 , and 4.79 ± 3.04 %bodyweight for the medio-lateral, antero-posterior, and vertical GRF components, respectively. Estimated force components were additionally strongly correlated with the ground truth measurements ($R^2 > 0.85$). Finally, we show high estimation accuracy for three clinically relevant point metrics on the paretic limb. These results suggest the potential for an individualized machine learning approach to translate to real-world clinical applications.

I. INTRODUCTION

Walking has numerous health and diagnostic benefits, and has been termed the “sixth vital sign” [1]. Neuromotor injuries such as stroke are leading causes of gait disability globally, with over 80% of survivors of stroke left with long-term locomotor impairments, a number that continues to grow as our population ages [2]. Post-stroke gait is characterized by hemiparesis, slow speeds, and asymmetry, largely due to reduced force production on the paretic, or more impaired, limb [3]. Thus, an overarching goal of gait rehabilitation for people post-stroke is to improve paretic limb force generation [4], [5].

Indeed, numerous clinical biomechanics studies have shown that physical therapy can help improve muscle strength and function in people post-stroke towards increasing paretic force generation [6]. These forces are quantified by the 3-dimensional ground reaction forces (3D GRFs) measured by specialized equipment such as force plates. Specifically, the medio-lateral (ML) GRF component is linked to

balance measures [7], the antero-posterior (AP) component is associated with walking speeds and symmetry [8], and the vertical GRF (VGRF) component is indicative of limb loading [9]. GRFs are also a necessary precursor to compute derivative biomechanical measures, such as joint torques and powers, which provide additional insights into post-stroke gait impairments [10]. However, while task-relevant gait rehabilitation, such as community-based rehabilitation [11], is important for long-term efficacy of an intervention [12], there is currently no accepted method for obtaining GRFs in the clinic or in real-world environments. Access to such measures would enable clinicians to provide personalized goals based on an individual’s rehabilitation progress.

Prior work has shown that wearable sensors may provide one opportunity for estimating 3D GRFs in healthy individuals, but is underexplored in clinical populations [13], [14]. A common approach for estimating GRFs is using pressure insoles in conjunction with data-driven, or machine learning, methods, given that they provide a direct measure of the foot-environment interface and represent the distribution of load across the foot [15]–[17]. While promising, these approaches are often insufficient for translation to people post-stroke as errors exceed the expected variability in this population [18]. One potential explanation for the lack of performance in prior methods is that they rely primarily on linear machine learning models, but normal loading information from pressure insoles may not be linearly related to GRF components. This nonlinearity may be due to factors such as shear forces along the ML and AP axes or shoe compliance characteristics. Nonlinear machine learning approaches have been shown to be effective in other areas of biomechanics and force estimation [19]–[21], and thus there is an opportunity to investigate their integration for 3D GRF estimation.

In this work, we assess the use of three simple machine-learning based algorithms, a convolutional neural network (CNN), a fully connected network (FCN), and linear regression (LR), to accurately estimate 3D GRFs in people post-stroke. We use a commercial pressure insole to obtain a high-resolution map of the pressure distribution across the foot during walking at varying speeds. We evaluate GRF estimation accuracy across the stance phase, while the foot is on the ground, for the different methods. Finally, we compare the performance of the algorithms for estimating key clinically relevant point metrics.

*These authors contributed equally to this work.

¹School of Engineering and Applied Sciences, Harvard University, Cambridge, MA

²College of Health and Rehabilitation Sciences, Sargent College, Boston University, Boston, MA

† Corresponding author walsh@seas.harvard.edu

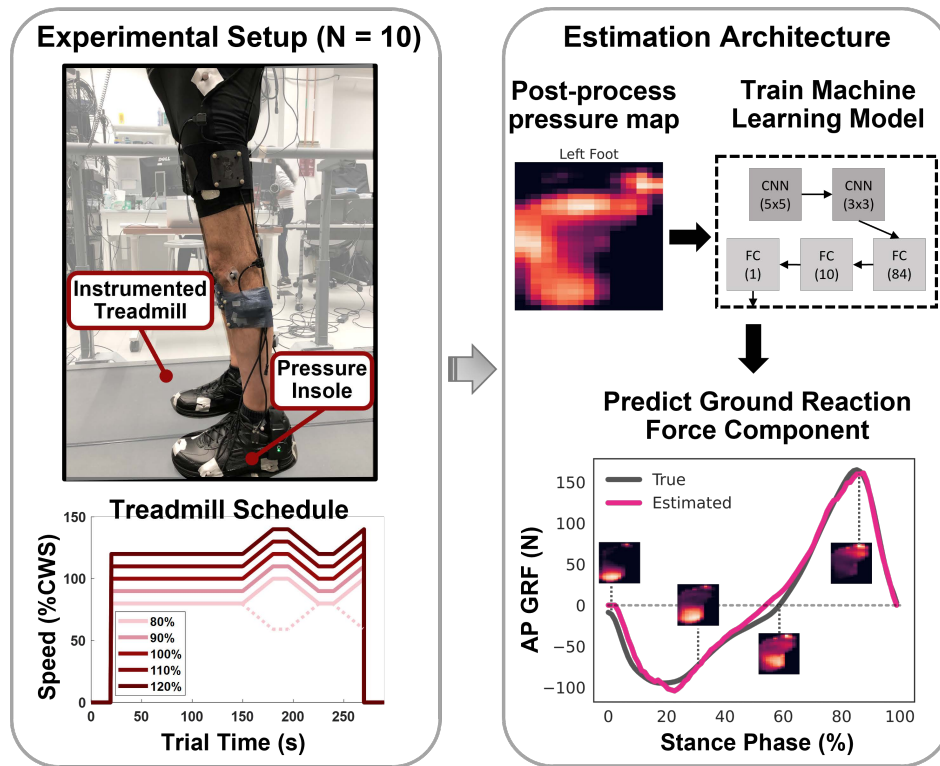


Fig. 1: (Left) Experimental setup and commanded treadmill speed throughout the 4-minute bouts. Treadmill speed increased first for half of the participants, and decreased first for the remaining participants (dashed profile). (Right) Pressure data from the insoles were reshaped into square heatmaps and used as input to train a simple machine learning algorithm to predict the corresponding 3D force vector component at each frame.

II. METHODS

A. Participants

Nine chronic post-stroke participants (>6 months post-stroke; 1F, 8M; age: 56 ± 13 yrs; height: 177 ± 6 cm; weight: 91 ± 20 kg) were recruited to participate in a single-session study. Six participants were left paretic, two participants wore a rigid ankle-foot-orthosis (AFO) on their paretic leg, and one participant completed two data collection sessions. The study was approved by the Harvard Longwood Medical Area Institutional Review Board and all individuals provided medical clearance and written informed consent.

B. Data Collection

Participants walked for a sequence of 3–6 4 min bouts on an instrumented split-belt treadmill (Bertec, Columbus, OH, USA; 2000 Hz), with the specific number of bouts depending on individual impairment and endurance levels (Table 1). Each bout was conducted at a different set of speeds defined by the participant’s comfortable walking speed (cws) and the order of bouts was randomized. Within each bout, the treadmill was commanded to remain at a constant speed for the first 2 min, and then ramped up and down through a range spanning 20%cws for the final 2 min (Treadmill Schedule in Figure 1). Half of the participants ramped up in speed first, while the remaining participants ramped down in speed first to account for any effects of acceleration versus

deceleration. We varied walking speeds to capture a wide spectrum of possible walking patterns for each individual and to diversify the dataset. A seated rest was provided between bouts. Force plates embedded in the treadmill measured 3D GRFs throughout the walking bouts. Participants wore commercial pressure insoles on both feet (XSensor, Calgary, AB, Canada; 50 Hz), each of which contained 233 individual pressure sensels. If the participant used an AFO, we placed the insole between the shoe and the underside of the AFO brace. Data obtained from the insole were time-synchronized with the force plates using a 5 V analog trigger signal sent at the onset of force plate recording. Participants were asked to step with each foot at the start and end of each walking bout, and corresponding peaks in pressure and VGRF data were used in post-processing to ensure data synchronization.

C. Data Processing

We then had two synchronized data sources for each trial of each participant, one from the insole and one from the instrumented treadmill. We applied a 2nd order low-pass Butterworth filter with a 10 Hz cutoff frequency on both sources of data to reduce noise artifacts. We also applied a minimum force threshold on the treadmill VGRF data to allow for robust gait event detection, such that all values less than 5% of the maximum VGRF in the trial were set to zero. We then normalized the insole data by subtracting the

average signal captured during the swing phase of the first stride in each trial, which we assumed should theoretically be zero in the absence of noise. Finally, we downsampled the treadmill data to 100 Hz and interpolated the insole data to match the treadmill data samples. We resized the insole sensel array from a 10×31 rectangular array, where elements corresponding to points outside the insole area did not map to a sensel, into a 28×28 square map using pixel area interpolation in OpenCV [22]. Note that while this approach distorts the aspect ratio of the pressure map, it does not alter the data represented. We used the GRFs from the treadmill to only include data from the stance phase. Then, the input data for the estimation models were the processed pressure data from each insole sensel during stance and the corresponding GRFs from the treadmill force plates were used as the ground truth to train and evaluate the models.

D. Model Development

Given the image form of the pressure maps, we used a CNN for this work (Figure 1). The network started with two convolutional layers with kernel sizes of 5 and 3. We used average pooling layers between the convolutional layers as we observed improved performance over max pooling. The convolutional layers were followed by two fully connected layers with 84 and 10 neurons. We added a final fully-connected layer to map pressure data from each foot to the GRF component at the corresponding frame. We used Rectified Linear Unit (ReLU) activation functions for the output of each layer prior to the final fully-connected layer to capture nonlinearity in the data. The model was trained using the Adam optimizer with a learning rate of 0.0005 and weight decay of 0.00001. These hyperparameters were empirically selected through a grid search. We used the mean squared error loss to train each model over 500 epochs. If the model failed to converge, we continued training until convergence or a maximum of 1000 epochs.

The model was trained individually for each foot of each participant, using at least one walking bout for training, one bout for validation, and one bout for testing. The slowest and fastest speeds were used for training, while the middle two speeds were used for validation and testing. Any remaining walking bouts were used as training data. If there were exactly three trials of usable data, then the slowest trial was used for training, the fastest speed was used for validation and the middle speed was used for testing. This train-test split allowed for generalizing the model across different speeds by accounting for the changes in gait biomechanics at varying speeds. One participant had less than three trials of usable data (due to too many strides with both feet on the same force plate and inconsistent insole connectivity), and was excluded from analysis. Table 1 lists the number of trials used for each participant.

To assess the efficacy of this approach, we compared the estimation accuracy of the CNN against an FCN and linear regression. The FCN had two hidden layers of 1024 neurons each, and each layer was fed through a ReLU activation function, making the model nonlinear. We used

the same training hyperparameters for the FCN models as in the CNN models. The linear regression provided a benchmark for conventional approaches in literature, while the FCN provided a benchmark for a nonlinear method that is not designed to capture spatial relationships. For the linear regression and FCN methods, the original rectangular insole data was reshaped into a 1D vector input. All CNN and FCN code was written using Pytorch 1.13.1 [23] and the linear regression code was written with Scikit-Learn 1.2.2 [24].

E. Model Evaluation

We evaluated the accuracy of the models by computing the root-mean square error (RMSE) between the estimated and true force at each frame in the test set for each individual and foot. We computed this error scaled to each individual's bodyweight to enable across-subject evaluation of the model. We also calculated the coefficient of determination, or R-squared, between the estimated and true force for each model. Finally, we assessed the RMSE for three key point metrics for the paretic limb that quantify impairment level: peak propulsion, propulsion impulse, and average VGRF in stance. Point metrics were calculated after segmenting the test set into stance cycles using VGRF data. Peak propulsion was quantified as the positive peak in AP GRF and propulsion impulse was determined as the positive integral of AP GRF for each stance cycle. Similarly, we computed the average VGRF for each stance cycle. We chose these metrics to enable comparison against documented minimal detectable change (MDC) thresholds that represent the expected within-day gait variability in this population [18]. If the estimation error for a participant with a single method differed from the other two methods by over two orders of magnitude, the result was considered an outlier and removed from further analysis. Finally, we used a linear mixed-effects model to assess the effect of estimation method on accuracy in which we defined subjects as the random-effects term, and the algorithm-type as the fixed-effects term. An alpha level of 0.05 was used to indicate significance.

III. RESULTS

A. Overall Estimation Accuracy

We observed that across all participants and across both feet, all three models were able to estimate 3D GRFs and were not statistically different from each other. After removing outliers, we found an RMSE of 0.88 ± 0.26 (mean \pm s.e.m.), 0.85 ± 0.27 , and 0.84 ± 0.25 %bodyweight (%bw) on the test datasets using the LR, FCN, and CNN-based methods respectively for paretic ML GRF, and 0.73 ± 0.19 , 0.71 ± 0.23 , and 0.66 ± 0.21 %bw on the non-paretic limb. For paretic AP GRF, we found an RMSE of 1.30 ± 0.66 , 1.28 ± 0.68 , and 1.21 ± 0.68 %bw, and for VGRF, 5.39 ± 4.34 , 5.98 ± 4.02 , and 5.63 ± 4.15 %bw using the LR, FCN, and CNN, respectively. On the non-paretic limb, we found errors in AP GRF estimates of 1.18 ± 0.44 , 1.07 ± 0.54 , and 1.05 ± 0.48 %bw, and in VGRF estimates of 4.02 ± 2.72 , 4.38 ± 2.89 , and 3.94 ± 2.78 %bw with the LR, FCN, and CNN, respectively. Figure 2 presents an

example of the true versus estimated GRF across a single stance cycle of a single participant with the median RMSE from the CNN approach. Table 1 presents the average performance across both limbs for each individual and method.

B. Correlation of Estimator versus Ground truth

In addition to estimation accuracy across stance, we evaluated the correlation between the estimated and true GRFs. After removal of outliers, we observed strong and statistically significant correlations ($R^2 > 0.85$, $p < 0.001$) for all three estimation approaches and all participants. Specifically, averaged across participants and both legs, we found an R^2 for ML GRF of 0.87 ± 0.11 , 0.86 ± 0.15 , and 0.87 ± 0.13 with the LR, FCN, and CNN, respectively. Similarly, we found R^2 values for AP GRF of 0.92 ± 0.08 , 0.91 ± 0.11 , and 0.91 ± 0.12 , and for VGRF of 0.95 ± 0.06 , 0.94 ± 0.07 , and 0.95 ± 0.06 using the LR, FCN, and CNN, respectively. A subset of the true versus estimated GRFs in the test dataset obtained with the CNN-based approach for all participants is shown in the bottom panel of Figure 2 and depicts the variance in estimator performance across the stride and across individuals. We observe that for most participants, there is close alignment between the true and estimated GRF across all points.

C. Point Metric Estimation Accuracy

Finally, we evaluated the performance of these approaches when estimating clinically relevant point metrics for the paretic limb. Similar to the estimator performance across the time series data, we found that the FCN and CNN-based approaches best estimated paretic propulsion metrics, but not statistically significantly. The average RMSE of peak paretic propulsion was 0.78 ± 0.39 %bw with the FCN and 0.82 ± 0.35 %bw with the CNN, close to the MDC of 0.8 %bw, versus 1.21 ± 0.55 %bw with the LR approach (Table 1). Average VGRF was best estimated by LR, with an error of 1.12 ± 0.50 %bw (MDC: 1.7 %bw), compared to 1.52 ± 1.32 %bw with the FCN and 1.34 ± 1.04 %bw with the CNN.

IV. DISCUSSION AND CONCLUSION

This work investigated the efficacy of three simple machine learning-based models to estimate 3D GRFs in people post-stroke using pressure insoles. We found that overall, a CNN was most effective, followed closely by the FCN, and lastly by linear regression. We further showed high estimation accuracy of key clinically-relevant point metrics, suggesting the potential for translation.

Similar to prior work in GRF estimation, we found a difference in performance across the different force components [16], [17]. VGRF estimates had the highest RMS errors, but the highest correlations. Conversely, AP GRF and ML GRF estimation errors were smaller, but correlations were less strong. The differences in estimation error reflect the varying magnitudes of the different GRF components. For example, the maximum amplitude of VGRF in our participants was approximately 10 times higher than that of AP GRF. The

differences in correlation measures reflect the hypothesized relationship between the sensor and estimated variable. The insoles measure pressure from normal loading distributed across the foot, and thus is most correlated with VGRF. We found that on average, performance as measured by error and correlation metrics was similar between the two nonlinear approaches, the CNN and FCN, both of which were consistently better than linear regression, although not statistically significantly. Anecdotally, we observed that the CNN converged faster during training than the FCN, which may be due to the ability of a CNN to better account for the spatial information contained in the input data [25]. Overall, these results suggest that when the input-output relationship is closer to linear, such as between the input pressure data and VGRF, performance is less sensitive to the choice of model. However, when there is strong nonlinear influence from factors such as shear forces and shoe compliance, as in the case of ML GRF and AP GRF, model selection becomes critical for estimation performance.

Interestingly, while on average, all three methods performed similarly on the time series data, the nonlinear methods provided an additional advantage over linear regression when computing key paretic propulsion metrics. Paretic propulsion metrics are often used to characterize the impairment level of an individual post-stroke [8] given its well-documented relationship with walking speed [4]. We find that with a CNN or FCN approach and the pressure insole, we obtain an average RMSE of 0.8 %bw in peak paretic propulsion compared to 1.2 %bw using linear regression. Furthermore, this improved performance of the nonlinear approaches over linear regression is observed consistently for nine out of ten datasets. Given that the MDC threshold for peak propulsion during single-session treadmill walking is 0.8 %bw, these methods may be able to differentiate between true changes in propulsion in this population and in turn, inform clinical assessments. Similarly, these estimation approaches achieve errors below the MDC for average paretic VGRF in stance, and thus may also be suitable for assessing limb loading characteristics during gait.

While these results are presented after removing outliers, we note that there is an additional risk of the standard linear regression used as a benchmark in this work. For one subject, the coefficients obtained through the linear regression least-squares cost minimization led to a coefficient on the order of 10^7 corresponding to a sensel that measured no pressure in the training dataset. However, a non-zero measurement by that sensel in the test dataset resulted in large errors for this individual and foot. Future work could investigate mitigating this issue by adding a regularization term to minimize the number of non-zero coefficients [26], albeit at the cost of losing some information from the input. Conversely, given the architecture of a CNN, which applies spatial filtering across groups of sensels, the risk of errors due to differences in the training versus testing distribution for a single sensel may be reduced, suggesting its potential for improved robustness.

There are a few limitations to note in this work that should

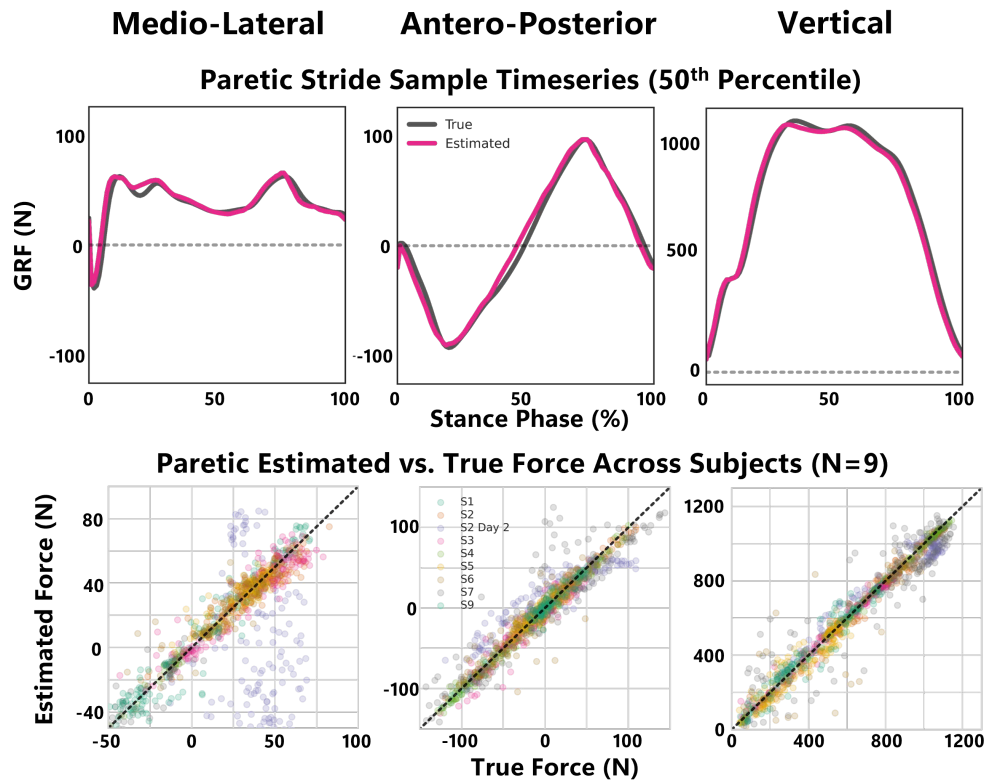


Fig. 2: (Top) CNN-estimated and force plate-based GRF time series data from a paretic stride in the test set of a single participant. Each plot represents the performance of the estimator with 50th percentile RMSE. (Bottom) CNN-estimated versus true force data from all participants' paretic limb test sets. The black dotted line represents perfect estimation.

			Test RMSE ¹ (%bw)									Paretic Point Metrics RMSE								
Worst ← RMSE → Best			ML GRF			AP GRF			V GRF			Peak Propulsion (%bw)			Propulsion Impulse (%bw %stance)			Average V GRF (%bw)		
Subj ID	CWS (m/s)	No. of trials	LR	FCN	CNN	LR	FCN	CNN	LR	FCN	CNN	LR	FCN	CNN	LR	FCN	CNN	LR	FCN	CNN
1	0.8	5	0.799	0.666	0.666	1.142	0.887	0.829	2.295	2.030	1.884	1.108	0.604	0.559	0.494	0.273	0.251	0.721	0.389	0.368
2	0.55	6	0.554	0.568	0.537	0.930	0.628	0.642	2.305	2.738	2.306	1.966	0.671	0.784	0.923	0.261	0.264	1.357	1.215	1.165
2 ²	0.55	5	0.553	0.505	0.545	0.748	0.752	0.708	3.032	2.875	2.834	0.541	0.622	0.468	0.282	0.339	0.252	0.510	0.492	0.551
3	0.9	6	1.058	0.987	1.020	1.719	1.694	1.230	4.533	6.029	4.763	1.389	1.718	1.187	0.835	0.644	0.604	1.237	4.443	1.032
4 ³	0.5	6	0.647	0.387	0.362	0.513	0.434	0.473	1.079	1.090	1.332	0.447	0.440	0.571	0.238	0.247	0.277	0.522	0.522	0.649
5 ⁴	0.4	6	0.912	1.021	1.030	1.544	1.679	1.733	9.446	8.934	8.469	-	1.010	1.056	-	0.541	0.531	-	2.210	2.645
6	0.7	5	0.941	1.031	0.963	1.823	1.921	1.773	9.320	9.675	9.479	0.987	0.727	0.912	0.537	0.343	0.403	1.975	1.544	1.893
7	1.0	6	0.931	0.982	0.863	1.957	1.757	1.852	7.942	8.873	7.584	1.425	0.721	1.404	0.700	0.513	0.676	1.289	2.261	3.265
8 ⁵	0.8	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9 ³	0.5	3	0.897	0.869	0.776	0.928	0.847	0.928	4.393	4.367	4.416	1.837	0.460	0.427	0.569	0.335	0.309	1.370	0.560	0.511
Average:			0.810	0.780	0.751	1.256	1.178	1.130	4.927	5.179	4.785	1.212	0.775	0.819	0.572	0.388	0.396	1.123	1.515	1.342
Std Dev:			0.184	0.250	0.238	0.518	0.574	0.539	3.191	3.302	3.039	0.552	0.391	0.345	0.242	0.142	0.166	0.504	1.319	1.035

Table 1. Individual-level estimator performance across models and metrics. ¹Averaged across both paretic and non-paretic limbs. ²Second visit for S2. ³Wore an ankle-foot-orthosis (AFO) on the paretic limb during the experiment. ⁴Paretic side excluded from LR analysis due to large outlier. ⁵Excluded due to lack of usable data.

be considered for future investigation. First, this study did not assess the ability of the model to predict performance across days. Given the known variability across days [18], more data may be needed to achieve similar prediction accuracy.

Another consequence of the single-session study design is that the insole placement was not adjusted between trials and thus, further investigation is needed to evaluate model performance with varying sensor placement. Second, while

this work was motivated by real-world applications, the use of this approach has yet to be tested in overground environments. As treadmill and overground walking are biomechanically different, further training may be required in order to accurately estimate forces in non-treadmill environments. Future work may investigate how to leverage data across individuals or across feet to increase the training dataset size and improve estimation accuracy. An additional approach may be to leverage temporal information rather than only spatial information, i.e., a time series of input frames rather than a single input frame [27]. Moreover, recent work has shown promising results using inertial measurement units, which capture kinematic information, to estimate AP GRF in people post-stroke during overground walking [11], [28]. Thus, future efforts may aim to develop methods that integrate kinematic and kinetic information from wearable sensors to better estimate 3D GRFs in varying environments.

In summary, this work demonstrates an approach for using pressure insoles to estimate 3D ground reaction forces in people post-stroke using machine learning-based nonlinear methods. While only with a few participants, this work presents initial work towards bridging the gap between clinical biomechanics and clinical rehabilitation through accurate estimation of kinetics in real-world environments.

ACKNOWLEDGMENT

The authors thank Harini Kannan, Ada Huang, Keysa Garcia, Dionna O'Roberts Williams, and our participants for their time and contributions to this study. This work was supported by funding from the National Institutes of Health (BRG-R01HD088619), the National Science Foundation (CMMI-1925085), and the Harvard University John A. Paulson School of Engineering and Applied Sciences.

REFERENCES

- [1] S. Fritz and M. Lusardi, "White paper: "walking speed: the sixth vital sign".." *Journal of geriatric physical therapy (2001)*, vol. 32, no. 2, pp. 46–9, 2009.
- [2] C. W. Tsao, A. W. Aday, Z. I. Almarzooq, A. Alonso, A. Z. Beaton, M. S. Bittencourt, A. K. Boehme, A. E. Buxton, A. P. Carson, Y. Commodo-Mensah, *et al.*, "Heart disease and stroke statistics—2022 update: a report from the american heart association," *Circulation*, vol. 145, no. 8, pp. e153–e639, 2022.
- [3] S. J. Olney and C. Richards, "Hemiparetic gait following stroke. Part I: Characteristics," *Gait & Posture*, vol. 4, no. 2, pp. 136–148, 1996.
- [4] L. N. Awad, M. D. Lewek, T. M. Kesar, J. R. Franz, and M. G. Bowden, "These legs were made for propulsion: advancing the diagnosis and treatment of post-stroke propulsion deficits," *Journal of NeuroEngineering and Rehabilitation*, vol. 17, p. 139, dec 2020.
- [5] J. F. Alingh, B. E. Groen, E. H. F. Van Asseldonk, A. C. H. Geurts, and V. Weerdesteyn, "Effectiveness of rehabilitation interventions to improve paretic propulsion in individuals with stroke - A systematic review," *Clinical biomechanics (Bristol, Avon)*, vol. 71, pp. 176–188, jan 2020.
- [6] J. J. Eng and P. Fang Tang, "Gait training strategies to optimize walking ability in people with stroke: A synthesis of the evidence," *Expert Rev Neurother*, vol. 7, no. 10, pp. 1417–1436, 2007.
- [7] R. R. Holt, D. Simpson, J. R. Jenner, S. G. Kirker, and A. M. Wing, "Ground reaction force after a sideways push as a measure of balance in recovery from stroke.," *Clinical rehabilitation*, vol. 14, pp. 88–95, feb 2000.
- [8] S. A. Roelker, M. G. Bowden, S. A. Kautz, and R. R. Neptune, "Paretic propulsion as a measure of walking performance and functional motor recovery post-stroke: A review," *Gait and Posture*, vol. 68, no. October 2018, pp. 6–14, 2019.
- [9] V. S. Mercer, J. K. Freburger, S.-H. Chang, and J. L. Purser, "Measurement of Paretic-Lower-Extremity Loading and Weight Transfer After Stroke," *Physical Therapy*, vol. 89, pp. 653–664, jul 2009.
- [10] T. H. Cruz, M. D. Lewek, and Y. Y. Dhaer, "Biomechanical impairments and gait adaptations post-stroke: Multi-factorial associations," *Journal of Biomechanics*, vol. 42, no. 11, pp. 1673–1677, 2009.
- [11] R. W. Nuckols, C. Chang, D. Kim, A. Eckert-Erdheim, D. Orzel, L. Baker, T. Baker, N. C. Wendel, B. Quinlivan, P. Murphy, J. Grupper, J. Villalobos, L. N. Awad, T. D. Ellis, and C. J. Walsh, "Design and evaluation of an independent 4-week, exosuit-assisted, post-stroke community walking program," *Annals of the New York Academy of Sciences*, may 2023.
- [12] M. M. Ardestani, C. E. Henderson, and T. G. Hornby, "Improved walking function in laboratory does not guarantee increased community walking in stroke survivors: Potential role of gait biomechanics," *Journal of Biomechanics*, vol. 91, pp. 151–159, 2019.
- [13] E. Shahabpoor and A. Pavic, "Measurement of Walking Ground Reactions in Real-Life Environments: A Systematic Review of Techniques and Technologies," *Sensors*, vol. 17, p. 2085, sep 2017.
- [14] W. Tao, T. Liu, R. Zheng, and H. Feng, "Gait Analysis Using Wearable Sensors," *Sensors*, vol. 12, pp. 2255–2283, feb 2012.
- [15] A. M. Howell, T. Kobayashi, H. A. Hayes, K. B. Foreman, and S. J. M. Bamberg, "Kinetic Gait Analysis Using a Low-Cost Insole," *IEEE Transactions on Biomedical Engineering*, vol. 60, pp. 3284–3290, dec 2013.
- [16] H. Rouhani, J. Favre, X. Crevoisier, and K. Aminian, "Ambulatory assessment of 3D ground reaction force using plantar pressure distribution," *Gait & Posture*, vol. 32, pp. 311–316, jul 2010.
- [17] D. T.-P. Fong, Y.-Y. Chan, Y. Hong, P. S.-H. Yung, K.-Y. Fung, and K.-M. Chan, "Estimating the complete ground reaction forces with pressure insoles in walking," *Journal of Biomechanics*, vol. 41, pp. 2597–2601, aug 2008.
- [18] T. M. Kesar, S. A. Binder-MacLeod, G. E. Hicks, and D. S. Reisman, "Minimal detectable change for gait variables collected during treadmill walking in individuals post-stroke," *Gait & Posture*, vol. 33, no. 2, pp. 314–317, 2011.
- [19] H. Prasanth, M. Caban, U. Keller, G. Courtine, A. Ijspeert, H. Vallery, and J. von Zitzewitz, "Wearable Sensor-Based Real-Time Gait Detection: A Systematic Review," *Sensors*, vol. 21, p. 2727, apr 2021.
- [20] B. B. Kang, D. Kim, H. Choi, U. Jeong, K. B. Kim, S. Jo, and K.-J. Cho, "Learning-Based Fingertip Force Estimation for Soft Wearable Hand Robot With Tendon-Sheath Mechanism," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 946–953, 2020.
- [21] E. Rho, D. Kim, H. Lee, and S. Jo, "Learning Fingertip Force to Grasp Deformable Objects for Soft Wearable Robotic Glove with TSM," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 8126–8133, 2021.
- [22] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.
- [23] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, pp. 8024–8035, Curran Associates, Inc., 2019.
- [24] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [25] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [26] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.
- [27] D. Kim, B. B. Kang, K. B. Kim, H. Choi, J. Ha, K.-J. Cho, and S. Jo, "Eyes are faster than hands: A soft wearable robot learns user intention from the egocentric view," *Science Robotics*, vol. 4, no. 26, p. eaav2949, 2019.
- [28] D. A. Revi, A. M. Alvarez, C. J. Walsh, S. M. De Rossi, and L. N. Awad, "Indirect measurement of anterior-posterior ground reaction forces using a minimal set of wearable inertial sensors: from healthy to hemiparetic walking," *Journal of NeuroEngineering and Rehabilitation*, vol. 17, p. 82, dec 2020.