Estimation of Ground Reaction Force Using Deep Neural Networks from Accelerometer Data: An Approach with Bi-LSTM, TCN, and Hybrid Architecture

Sérgio Rodrigues Lima Jr, Ronaldo F. Zampolo and Antônio Pereira Jr.

***Abstract*—This study presents a deep learning approach to estimate ground reaction force from accelerometer data using Bi-LSTM, TCN, and a hybrid architecture. A cross-correlation analysis was performed to identify the sensor with the most informative signals for prediction. The hybrid model achieved the best balance between accuracy and training time, showing promising results in RMSE, rRMSE, and R^2. The proposed methodology demonstrates potential for real-time gait analysis in wearable systems, offering a portable and low-cost alternative for clinical and sports applications.**

***Keywords*- Ground Reaction Force, Inertial Sensors, Gait Analysis, Deep Learning, Wearable Devices**

I. INTRODUCTION

Ground Reaction Force (GRF) is a fundamental biomechanical variable in human motion analysis, frequently employed in gait studies, clinical evaluations, and sports monitoring. However, its direct measurement requires the use of force plates (with costs around $3,437 [1]) or instrumented insoles, which limits its applicability to controlled environments [2].

Inertial sensors have emerged as a portable and low-cost alternative for indirect GRF estimation in out-of-laboratory contexts, promoting greater technological inclusion in health and sports applications [3], [4].

The relationship between accelerometer signals and GRF has been widely investigated, although it presents significant modeling challenges due to its nonlinear and time-dependent nature. Small variations in acceleration can generate disproportionate changes in the estimated force, and this relationship varies throughout the gait cycle, since the forces generated in each phase (such as initial contact, loading response, midstance, and push-off) exhibit distinct biomechanical characteristics. This complexity demands advanced modeling techniques capable of capturing both the nonlinear variations and the dynamic changes over time [5]. In this context, deep neural networks such as Bi-LSTM (Bidirectional Long Short-Term Memory) and TCN (Temporal Convolutional Networks) have shown great potential for analyzing complex time series [6], [7].

Bi-LSTM networks are an extension of conventional LSTMs, capable of capturing temporal dependencies in both directions — past and future — making them particularly suitable for tasks involving contextual analysis over time [3]. On the other hand, TCNs use dilated causal convolutions to efficiently model temporal sequences, preserving the temporal order of the data and enabling long-term pattern learning with lower computational cost compared to recurrent networks [4], since causality ensures that predictions at time t do not depend on future inputs, while dilation expands the receptive field by spacing out the elements considered in the convolution, allowing the capture of long-range dependencies with fewer layers [7].

This work proposes a comparison between three neural network architectures for GRF prediction based on accelerometer signals: one based on Bi-LSTM, another on TCN, and a third hybrid model combining both. The study uses the dataset described in [8], which allows for the comparison, validation, and improvement of different motion capture systems. We used accelerometer recordings from inertial sensors placed on the lumbar region, thigh, and foot, as well as GRF data obtained from instrumented insoles on the right foot. A cross-correlation analysis was performed to identify the sensor most correlated with the GRF, and the models were evaluated using the RMSE, rRMSE, and R^2 metrics. The results demonstrate the potential of the proposed networks for the development of portable and low-cost systems for gait monitoring, enabling clinical and sports applications outside the laboratory environment.

At the following link is the repository containing the training codes for each neural network, the calculation of the mean cross-correlation for each sensor, and the GRF prediction:

`github.com/Network-Trainning.git`

II. METHODOLOGY

A. Dataset

The dataset used in this study was originally described by [8] and includes detailed biomechanical recordings of human gait acquired simultaneously from multiple systems. Among these, we selected synchronized signals from inertial measurement units (IMUs) and force insoles. Data were collected from ten asymptomatic participants performing several locomotor tasks in a laboratory setting. Specifically, this study used trials of normal gait, slow gait, fast gait, and two-minute walking.

Sérgio Rodrigues Lima Jr, Ronaldo de Freitas Zampolo, and Antônio Pereira Jr. are with the Signal Processing Laboratory (LaPS), Federal University of Pará (UFPA), Belém-PA, Brazil. Emails: {serginholimaaudio@gmail.com, zampolo@ufpa.br, apereira@ufpa.br.

Accelerometer data were acquired from Physilog®6S sensors, placed on eight anatomical locations (lumbar, pelvis, thighs, shanks, and feet), capturing tri-axial linear acceleration, angular velocity, and magnetic field at 256 Hz. In this work, we used only the tri-axial acceleration data from the thighs, shanks, and feet. GRF data were recorded at 100 Hz using Insole3 instrumented insoles, equipped with 16 pressure sensors and an embedded inertial system to estimate total force and center of pressure. All signals were obtained from synchronized `.csv` files provided by the dataset authors, aligned using predefined gait events such as acceleration peaks and vertical jumps.

B. Data Preprocessing

In order to match the dynamic range of different sensors and reduce the influence of noise, normalization and filtering procedures were applied to both the input and output data during neural network training. Z-score normalization was applied to the input data (accelerometers) to rescale the variables, reducing the impact of outliers and improving convergence during model training [9]–[11].

The output variable (GRF) was also normalized. This additional normalization aimed to balance the scale between input and output, avoiding the disproportionate influence of the GRF's high magnitude on the loss function [12].

Filtering was performed using a fourth-order Butterworth low-pass filter with a cutoff frequency of 10 Hz, following the same configuration described in [8], with the goal of eliminating high-frequency noise and preserving the relevant components of biomechanical signals [13]–[15].

C. Cross-correlation analysis between accelerometer signals and GRF

In order to identify which inertial sensor provides the most informative data for GRF prediction, a cross-correlation analysis was performed between the resultant accelerometer signals and the GRF. This method allows quantifying the similarity between two time series, considering potential time lags between them [16].

Given two discrete signals, $x[n]$ (accelerometer signal) and $y[n]$ (GRF signal), the cross-correlation function $R_{xy}[\tau]$ is defined as:

$$R_{xy}[\tau] = \sum_{n} x[n]y[n+\tau], \tag{1}$$

where τ represents the lag (positive or negative). This function evaluates how much one signal resembles the other as it is shifted in time.

The resultant acceleration magnitude a_{res} was calculated using the following equation:

$$a_{\text{res}} = \sqrt{a_x^2 + a_y^2 + a_z^2}, \tag{2}$$

where a_x, a_y, and a_z correspond to the accelerations recorded along the three orthogonal axes.

The normalized cross-correlation was calculated between the resultant acceleration signal and the GRF for each recording. This normalization ensures that the correlation values are bounded between -1 and 1, analogous to computing the Pearson correlation coefficient for each lag [16]. By normalizing, the influence of signal amplitude is removed, allowing the comparison to focus solely on the shape similarity between the signals.

From the cross-correlation function, the maximum absolute correlation value was extracted, representing the degree of similarity between the signals, regardless of the lag at which it occurred. This procedure was repeated for all files in the dataset.

Finally, for each sensor location, the mean and standard deviation of the correlation values were computed. These results, presented in Table I and Figure 1, allowed the identification of the sensor most correlated with the GRF, providing support for selecting the optimal location for GRF prediction models.

Table I
MEAN AND STANDARD DEVIATION OF THE CROSS-CORRELATION BETWEEN ACCELEROMETER AND GRF DATA

Sensor	Mean Correlation	Standard Deviation
Lumbar	0.7841	0.0716
Thigh	0.7799	0.0708
Shank	0.7854	0.0711
Foot	**0.8521**	**0.0802**

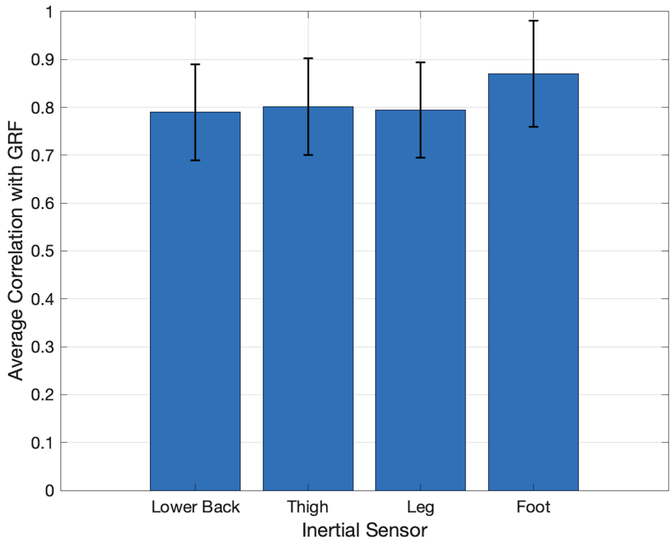

Figure 1. Average cross-correlation between the inertial sensor signals and the GRF. Bar height represents the mean correlation for each sensor, while the error bars indicate the standard deviation of correlation values across different files analyzed.

D. Network Architecture Details

The configurations of the Bi-LSTM, TCN, and Hybrid models were designed to balance predictive performance and computational efficiency. All models were implemented using MATLAB's Deep Learning Toolbox and trained with the Adam optimizer for 30 epochs using a mini-batch size of 128 samples.

Bi-LSTM Model: Based on bidirectional long short-term memory, capable of capturing temporal dependencies in both directions [6], [18]. The architecture of the Bi-LSTM model had the following layers:

- Sequence input layer with 3 input channels corresponding to the accelerometer axes;
- Bi-LSTM layer with 128 hidden units;
- Dropout layer with dropout rate of 0.4;
- Bi-LSTM layer with 64 hidden units;
- Dropout layer (rate = 0.4);
- Fully connected layer with 64 units followed by a ReLU activation;
- Fully connected layer with 1 output neuron;
- Regression layer using the mean squared error loss.

This model has approximately 155,000 trainable parameters, with the majority concentrated in the recurrent layers.

TCN Model: A convolutional model with causal dilated convolutions, which processes sequences in parallel with a lower computational cost [7], [19]. The architecture of our TCN model included:

- Sequence input layer with 3 channels;
- Causal 1D convolutional layer with 128 filters, kernel size of 3, and dilation factor of 1;
- ReLU activation followed by a dropout layer (rate = 0.4);
- Causal 1D convolutional layer with 64 filters, kernel size of 3, and dilation factor of 2;
- ReLU activation and dropout (0.4);
- Global max pooling layer;
- Fully connected layer with 64 units and ReLU activation;
- Fully connected layer with 1 neuron, followed by a regression output layer.

The model contains approximately 85,000 trainable parameters.

Hybrid Model (TCN + Bi-LSTM): A combination of TCN convolutional layers with Bi-LSTM layers, aiming to extract local and temporal patterns in a complementary manner. The hybrid architecture combines convolutional and recurrent structures:

- Two causal convolutional layers:
  - First: 64 filters, kernel size 3, dilation factor 1;
  - Second: 128 filters, kernel size 3, dilation factor 2.
- Each convolutional layer is followed by batch normalization and ReLU activation;
- Two Bi-LSTM layers: first with 128 hidden units and second with 64 hidden unit;
- Final dense regression head:
  - Fully connected layer (64 units) $\rightarrow$ ReLU $\rightarrow$ Fully connected (1 unit) $\rightarrow$ Regression layer.

This architecture comprises approximately 195,000 trainable parameters.

E. Time Windowing and Validation

The time windowing step aims to transform the continuous accelerometer signals into input sequences suitable for training neural networks. To achieve this, the script scans the data using sliding windows of fixed length (30 samples) with strides of 1, 3, and 5 samples, generating a set of temporal segments that represent successive portions of the movement. Each window contains a matrix with the components of the three-axis accelerometer signals over time, forming a three-dimensional input for the model. The output associated with each window is the GRF value corresponding to the last sample in the window, allowing the model to learn to predict the GRF based on the recent history of movement. This approach is essential for capturing the temporal dynamics of gait and providing sufficient context for the model to identify relevant patterns in the evolution of the signals. This segmentation strategy enables the capture of fine-grained local variations and transitional dynamics between gait phases, as detailed in [20].

For model evaluation, a 5-fold cross-validation (K=5) was employed. This approach enhances statistical robustness and maximizes the use of the available dataset, as supported by the methodologies outlined in [21], [22].

F. Model Training and Evaluation

All models were trained using the Adam optimizer [23], with the mean squared error (MSE) serving as the loss function. Performance evaluation was measured with the root mean squared error (RMSE), relative RMSE (rRMSE), and the coefficient of determination (R^2) as key metrics [24], [25]. Table II summarizes the average performance values obtained across the models.

Table II
MODEL PERFORMANCE FOR GRF PREDICTION

Model	**RMSE**	**rRMSE (%)**	R^2
Bi-LSTM	0.3005	6.32	0.9112
TCN	0.4958	10.64	0.7535
Hybrid	0.2931	6.10	0.9134

The inference time analysis of each model was not documented, as no significant differences were observed among them.

III. RESULTS AND DISCUSSION

A. Model Performance

Table II summarizes the results of the three architectures for GRF estimation. The hybrid model demonstrated the best overall performance and the highest coefficient of determination. The Bi-LSTM also performed well, while the TCN, despite being the most computationally efficient, exhibited the highest errors.

These results highlight the advantage of combining convolutional and recurrent networks to capture both short- and long-term temporal patterns in gait data, as previously explored in the literature [6], [7], [18].

B. Qualitative Analysis

Figure 3 presents a visual comparison between the actual GRF values and the predictions generated by the Bi-LSTM,

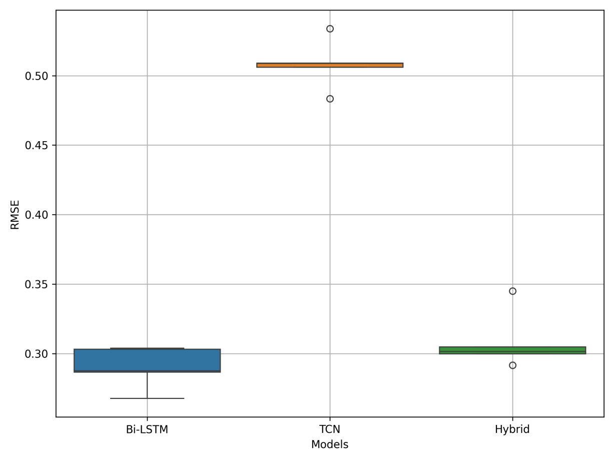

Figure 2. Boxplot of the RMSE values obtained in the 5 cross-validation folds for the Bi-LSTM, TCN, and hybrid models. It can be observed that the Bi-LSTM model presents the lowest average RMSE, indicating higher accuracy, while the TCN model shows the highest errors, despite its consistency across folds. The hybrid model combines good accuracy with low variability, suggesting a better balance between performance and robustness.

TCN, and Hybrid models. It can be observed that the Bi-LSTM model adequately follows the peaks and valleys of the GRF curve, while the TCN model tends to smooth the peaks and underestimate the maximum values. The Hybrid model, in turn, demonstrates greater fidelity to the curve shape, capturing the temporal variations of the GRF throughout the gait cycle more accurately, reinforcing its superiority in adherence to the actual signal.

C. Training Time

Table III shows the average training time per fold. TCN was the most efficient, while Bi-LSTM, the most accurate, had the highest computational cost. The hybrid model balanced performance and training time.

Table III
AVERAGE TRAINING TIME PER FOLD

Model	Average Time
Bi-LSTM	16 min 88 s
TCN	56 s
Hybrid	14 min 24 s

IV. CONCLUSION

This study proposed and evaluated a methodology for estimating GRF using deep learning models trained on accelerometer data from inertial sensors. Three distinct neural network architectures were investigated: Bi-LSTM, TCN, and a hybrid model combining both approaches. To ensure the quality of the signals used in training and evaluation, data normalization, low-pass filtering, and temporal segmentation techniques were applied.

A cross-correlation analysis was conducted to identify the sensor with the highest predictive potential, revealing that the accelerometer positioned on the foot showed the strongest correlation with GRF signals, supporting the choice of input data used in the three models.

Quantitative results demonstrated the effectiveness of the proposed architectures: the Bi-LSTM model achieved the highest accuracy, the TCN stood out for its computational efficiency, and the hybrid model achieved the best balance between performance and training time. Metrics such as RMSE, rRMSE, and R^2 confirmed the ability of the networks to accurately capture the dynamic variations of GRF throughout the gait cycle.

The ability to estimate GRF based on accelerometer data and deep learning techniques opens up relevant pathways for large-scale clinical and social applications. Clinically, this approach can enable low-cost and non-invasive gait analysis in non-laboratory or resource-limited environments, allowing continuous monitoring of patients with neurological or musculoskeletal conditions. Additionally, the portability and accessibility of inertial sensors may support large-scale screening and the detection of mobility impairments in elderly populations, contributing to preventive healthcare strategies. From a social perspective, integrating this technology into wearable devices can provide individuals with feedback on gait quality, encouraging their engagement in rehabilitation and promoting greater autonomy during physical recovery. Thus, the use of artificial intelligence-based models for GRF prediction represents a step forward toward more accessible, personalized, and biomechanically-informed health monitoring solutions.

As possible future studies, expanding the dataset with diverse populations, integrating attention mechanisms and Transformer-based models, developing domain adaptation strategies, and implementing the models in embedded real-time systems could be explored. Investigating gait symmetry and the models' generalization capacity for the left limb also represent promising directions.

In summary, this work provides a robust and accurate solution for estimating GRF data by combining biomechanical knowledge with the power of deep neural networks.

REFERENCES

[1] Gait and Motion Technology. *EasyBase Force Plates*. Disponível em: https://www.gaitandmotion.co.uk/product-page/easybase-force-plates. Acesso em: 21 abr. 2025.

[2] W. Tao, T. Liu, R. Zheng e H. Feng, "Gait analysis using wearable sensors," *Sensors*, vol. 12, no. 2, pp. 2255–2283, 2012.

[3] Jiang, X., Napier, C., Hannigan, B., Eng, J. J., & Menon, C. (2020). Estimating Vertical Ground Reaction Force during Walking Using a Single Inertial Sensor. Sensors, 20(15), 4345. https://doi.org/10.3390/s20154345

[4] Scheltinga, B. L., Kok, J. N., Buurke, J. H., & Reenalda, J. (2023). Estimating 3D ground reaction forces in running using three inertial measurement units. Frontiers in Sports and Active Living, 5, 1176466. https://doi.org/10.3389/fspor.2023.1176466

[5] G. Leporace, L. A. Batista, and J. Nadal, "Prediction of 3D ground reaction forces during gait based on accelerometer data," *Research on Biomedical Engineering*, vol. 34, no. 3, pp. 211–216, 2018. Available: https://doi.org/10.1590/2446-4740.06817

[6] P. Zhang, Y. Yang, and Z.-Y. Yin, "BiLSTM-Based Soil–Structure Interface Modeling," *International Journal of Geomechanics*, vol. 21, no. 7, p. 04021096, 2021. Available: https://doi.org/10.1061/(ASCE)GM.1943-5622.0002058

[7] S. Bai, J. Z. Kolter, and V. Koltun, "An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling," *arXiv preprint arXiv:1803.01271*, 2018.

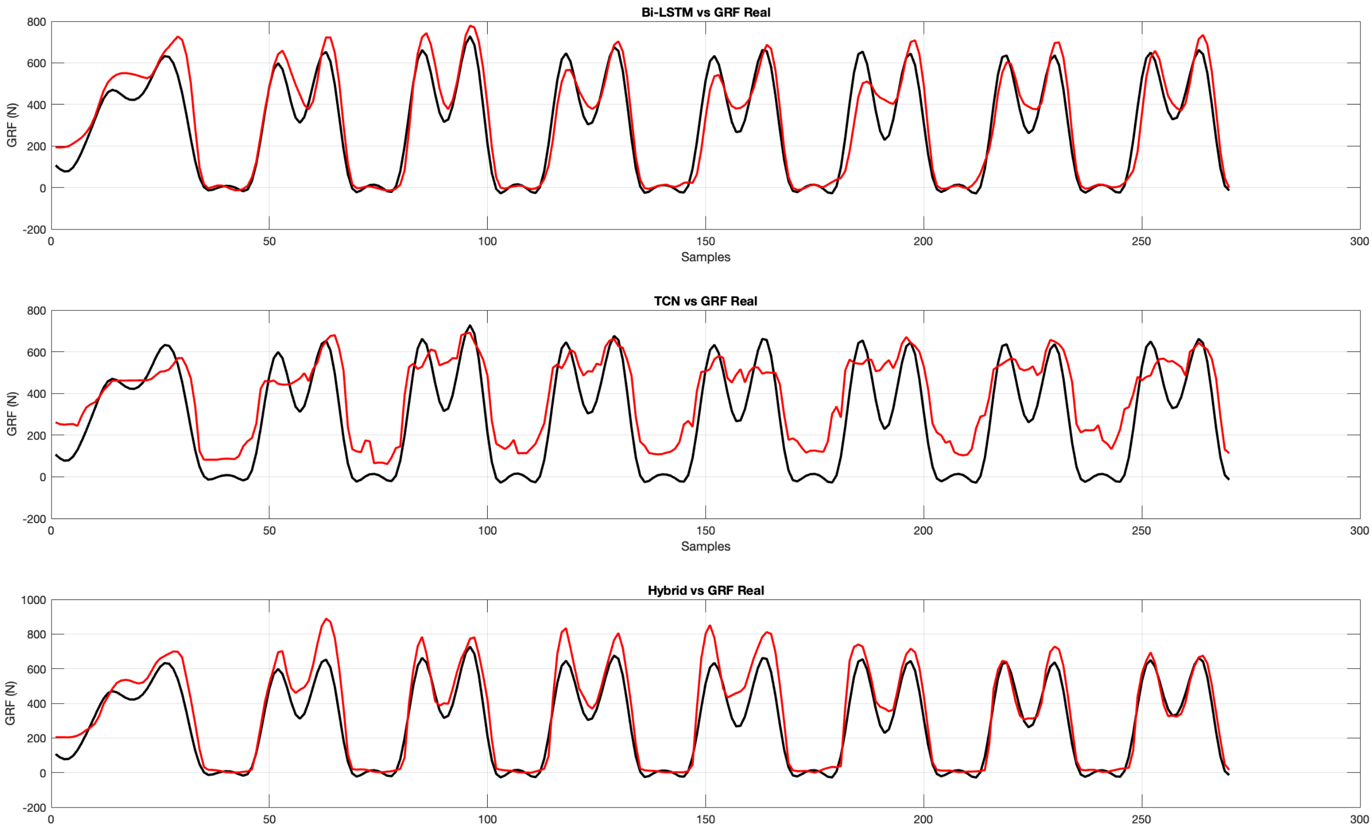

Figure 3. Comparison between the actual GRF and the predicted values obtained from three different neural network architectures: Bi-LSTM, TCN, and Hybrid (TCN + Bi-LSTM). Each subplot shows the actual GRF (black line) and the corresponding model estimate (red line). The visual comparison highlights the ability of each model to follow the temporal dynamics of the GRF signal throughout the gait cycle.

[8] Grouvel, G., Carcreff, L., Moissenet, F., & Armand, S. (2023). A dataset of asymptomatic human gait and movements obtained from markers, IMUs, insoles and force plates. *Scientific Data*, 10, 180. https://doi.org/10.1038/s41597-023-02077-3

[9] M. Z. Al-Faiz, A. A. Ibrahim, and S. M. Hadi, "The effect of Z-Score standardization (normalization) on binary input due to the speed of learning in back-propagation neural network," *International Journal of Information and Communication Technology*, vol. 1, no. 3, pp. 42–48, 2019. Available: https://doi.org/10.31987/IJICT.1.3.41

[10] H. Suprajitno, "Investigations on Impact of Feature Normalization Techniques for Prediction of Hydro-Climatology Data Using Neural Network Backpropagation with Three Layer Hidden," *International Journal of Sustainable Development and Planning*, vol. 17, no. 7, pp. 2069–2074, 2022. Available: https://doi.org/10.18280/ijsdp.170707

[11] J. Yun, H. Kim, S. Cho, and H. Kang, "ZNorm: Z-Score Gradient Normalization for Accelerating Neural Network Training," *arXiv preprint arXiv:2408.01215*, [Online]. Available: https://doi.org/10.48550/arxiv.2408.01215

[12] Y. LeCun, L. Bottou, G. B. Orr, and K. R. Müller, *Efficient BackProp*, in *Neural Networks: Tricks of the Trade*, Lecture Notes in Computer Science, vol. 7700, Springer, pp. 9–48, 2012.

[13] S. W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*, California Technical Publishing, 1997.

[14] S. K. Mitra, *Digital Signal Processing: A Computer-Based Approach*, 4th ed., McGraw-Hill, 2011.

[15] D. A. Winter, *Biomechanics and Motor Control of Human Movement*, Hoboken, NJ: John Wiley & Sons, 2009.

[16] G. C. Carter, "Coherence and time delay estimation," *Proceedings of the IEEE*, vol. 75, no. 2, pp. 236–255, 1987.

[17] A. Cutti *et al.*, "Outwalk: A protocol for clinical gait analysis based on inertial and magnetic sensors," *Med. Biol. Eng. Comput.*, vol. 48, no. 1, pp. 17–25, 2010.

[18] A. Graves, "Generating Sequences With Recurrent Neural Networks," *arXiv preprint arXiv:1308.0850*, 2013.

[19] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal Convolutional Networks for Action Segmentation and Detection," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[20] J. Taborri, E. Palermo, S. Rossi, and P. Cappa, "Gait Partitioning Methods: A Systematic Review," *Sensors*, vol. 16, no. 1, p. 66, 2016. Available: https://www.mdpi.com/1424-8220/16/1/66

[21] Y. Bengio and Y. Grandvalet, "No unbiased estimator of the variance of k-fold cross-validation," *Journal of Machine Learning Research*, vol. 5, pp. 1089–1105, 2004.

[22] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1137–1143, 1995.

[23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. International Conference on Learning Representations (ICLR)*, 2015. [Online]. Available: https://arxiv.org/abs/1412.6980

[24] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geoscientific Model Development*, vol. 7, pp. 1247–1250, 2014. Available: https://doi.org/10.5194/gmd-7-1247-2014

[25] D. Zhang, "A Coefficient of Determination for Generalized Linear Models," *The American Statistician*, vol. 71, no. 4, pp. 310–316, 2017. Available: https://doi.org/10.1080/00031305.2016.1256839