

## Article

# Prediction of Three-Directional Ground Reaction Forces during Walking Using a Shoe Sole Sensor System and Machine Learning

Takeshi Yamaguchi <sup>1,2,\*</sup> , Yuya Takahashi <sup>1</sup> and Yoshihiro Sasaki <sup>3</sup>

<sup>1</sup> Department of Finemechanics, Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan; yuya.takahashi.p1@dc.tohoku.ac.jp

<sup>2</sup> Graduate School of Biomedical Engineering, Tohoku University, Sendai 980-8579, Japan

<sup>3</sup> Research Institute for Electromagnetic Materials, Tomiya 981-3341, Japan; sasaki@denjiken.ne.jp

\* Correspondence: takeshi.yamaguchi.c8@tohoku.ac.jp

**Abstract:** We developed a shoe sole sensor system with four high-capacity, compact triaxial force sensors using a nitrogen added chromium strain-sensitive thin film mounted on the sole of a shoe. Walking experiments were performed, including straight walking and turning (side-step and cross-step turning), in six healthy young male participants and two healthy young female participants wearing the sole sensor system. A regression model to predict three-directional ground reaction forces (GRFs) from force sensor outputs was created using multiple linear regression and Gaussian process regression (GPR). The predicted GRF values were compared with the GRF values measured with a force plate. In the model trained on data from the straight walking and turning trials, the percent root-mean-square error (%RMSE) for predicting the GRFs in the anteroposterior and vertical directions was less than 15%, except for the GRF in the mediolateral direction. The model trained separately for straight walking, side-step turning, and cross-step turning showed a %RMSE of less than 15% in all directions in the GPR model, which is considered accurate for practical use.

**Keywords:** ground reaction force; shoe sole sensor system; machine learning; gait; walking; turning



**Citation:** Yamaguchi, T.; Takahashi, Y.; Sasaki, Y. Prediction of Three-Directional Ground Reaction Forces during Walking Using a Shoe Sole Sensor System and Machine Learning. *Sensors* **2023**, *23*, 8985. <https://doi.org/10.3390/s23218985>

Academic Editor: Giovanni Saggio

Received: 17 October 2023

Revised: 1 November 2023

Accepted: 2 November 2023

Published: 5 November 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the advent of aging in society, the number of falling accidents among elderly people and patients with gait disorders is increasing. Thus, fall prediction and rehabilitation based on balance assessment are becoming increasingly important. The measurement of ground reaction forces (GRFs) during gait is used in a wide range of medical and healthcare applications, including the assessment of fall risk [1], diagnosis of pathological gait [2], and biofeedback for rehabilitation [3]. Three-directional GRFs are usually measured using force plates because of their high accuracy. However, force plates cannot be easily introduced in small clinical facilities owing to their high cost, and space and movement limitations. Furthermore, it is difficult to measure GRFs with force plates during natural gait in a living environment.

Many attempts have been made to measure GRFs without using force plates. Studies have been conducted to estimate GRFs using kinematic data obtained from marker-less motion capture [4,5]. Liu et al. [6] found that if the cross-correlation coefficient is greater than 0.9 and the percent root-mean-square error (%RMSE) is less than 15%, the accuracy of GRF estimation is excellent. However, limited reports using marker-less motion capture met these criteria for GRF estimation in three-axis directions. Another problem is that the use of a camera also imposes location restrictions. GRF estimation using wearable sensors equipped with inertial measurement units (IMUs) is widely adopted because of its ease of use and low cost, and it is less subject to location and motion restrictions [7–9]. One potential problem with an IMU system is that its performance can be reduced by placement

error [10]. Shoe sole sensor systems incorporating devices that measure forces in shoes, such as pressure sensors embedded in insoles and triaxial force sensors attached to the outer sole of shoes, have been developed. Moreover, various types of low-cost in-shoe mobile measurement systems using a small number of force sensors have been developed in previous studies [11–15]. Furthermore, some groups have proposed the estimation of GRFs during walking from these local measurements, using machine learning techniques, such as linear [16–18] and nonlinear [19–21] regressions. The insole-type sensor uses only vertical forces to estimate three-directional GRFs, and a three-axis force sensor would enable estimation with the addition of horizontal (anteroposterior and mediolateral) force data, which could increase accuracy. Therefore, shoe sole sensor systems with three-axis force sensors attached to the outer sole of the shoe have been developed [22,23]. Although these systems are superior to previous systems as they can directly measure GRFs and have a small error rate, the possibility that they may affect natural gait cannot be denied. Moriyasu et al. [24] and Yamaguchi [25] developed sole sensor systems with a number of small triaxial force sensors mounted on the outer sole of a shoe to measure GRF distribution during running and walking, respectively, but the systems had problems, such as an inability to measure the vertical GRF component at the heel, where a high load is applied owing to the limited rated capacity of the sensors.

In the present study, we developed a lightweight sole sensor system with a high-capacity, compact triaxial force sensor using a nitrogen added chromium (Cr–N) strain-sensitive thin film [26,27] mounted at four points (heel, fifth metatarsal, first metatarsal, and toe) on the sole of a shoe, and estimated three-directional GRFs using machine learning. We aimed to investigate whether it is possible to estimate three-directional GRFs using machine learning, such as multiple linear regression (MLR) analysis and Gaussian process regression (GPR), with a small number of data measured using the sole sensor system. In addition, since it is difficult to obtain a large amount of data when considering the application to clinics, we used a GPR model [28–31] and cross-validation to enable estimation without overfitting even with a small amount of data.

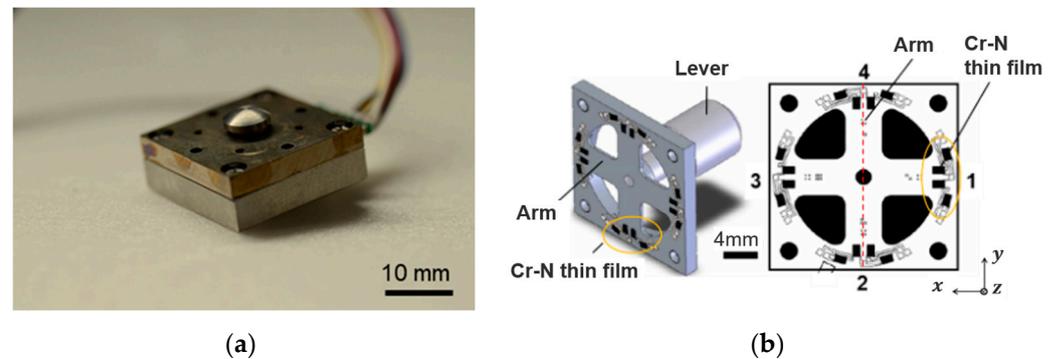
## 2. Methods

### 2.1. Sole Sensor System Using a Cr–N Strain-Sensitive Thin Film

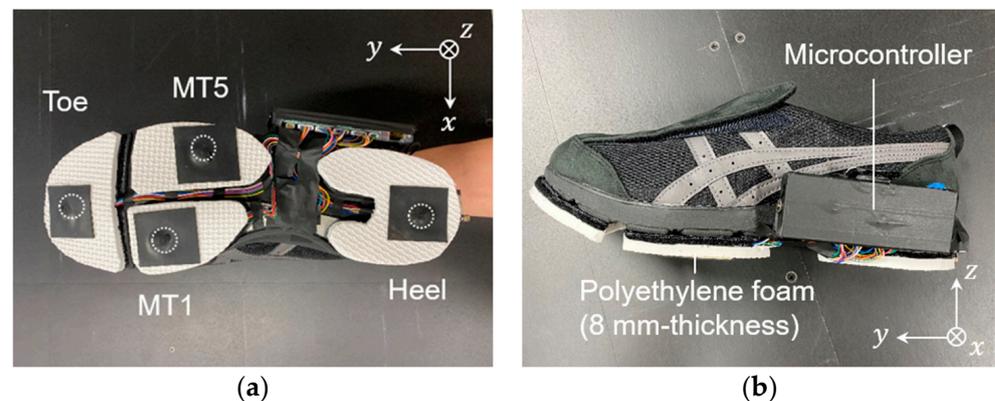
The triaxial force sensor using a Cr–N thin film (Research Institute for Electromagnetic Materials, Tomiya, Japan; dimensions: 20 mm × 20 mm × 7.5 mm; mass: 18 g) used in this study is shown in Figure 1. The sensor consists of a 20 mm square stainless-steel housing and a force-sensing contactor (lever). The rated capacities of the sensor in the  $x$ ,  $y$ , and  $z$  directions were  $\pm 500$  N,  $\pm 500$  N, and 1000 N, respectively. As shown in Figure 1b, the force sensor was fabricated by forming an insulating film on the stainless-steel strain structure and directly forming a Cr–N thin film on it through sputtering [26]. When a force acts on the tip of the contactor, voltage output signals are obtained from four sets of Cr–N strain gauge thin-film elements on the outer edge of each of the four arms, and the magnitude and direction of the applied force can be estimated by analyzing these signals [26,27].

Figure 2 shows the appearance of the shoe sole sensor system developed in this study. An 8 mm thick polyethylene foam outsole was attached to the sole of a walking shoe (LifeWalker Men’s FLC101, size: 27.0 cm; LifeWalker Women’s FLC307, size: 25.0 cm; ASICS, Kobe, Japan). A Cr–N thin-film strain sensor was attached to the partially cut-out portion of the sole. A total of eight sensors, four for each shoe, were mounted simultaneously to measure three-directional forces at each location. The sensors were covered with 1 mm thick nitrile rubber to prevent abrasion between the contactor of the sensors and the ground surface. The force sensors were wired to a microcontroller (Teensy 3.6, SparkFun Electronics<sup>®</sup>, Niwot, CO, USA), and a case containing the board and battery was attached to the side of the shoe. The mass of the entire shoe sole sensor system, including each sensor, battery, and board, was less than 380 g. Force data were recorded to an SD card for each measurement. The sampling frequency for data measurement was 870 Hz. As shown in Figure 2a, the  $x$ ,  $y$ , and  $z$  directions of the force measured by the small

triaxial force sensor were the foot width, foot length, and vertical direction of the shoe. The three-directional forces measured by the triaxial sensor were denoted as  $f_{xi}$ ,  $f_{yi}$ , and  $f_{zi}$ , respectively, where  $i$  denotes the position of the sensor,  $i = 1$  is the heel,  $i = 2$  is the first metatarsal,  $i = 3$  is the fifth metatarsal, and  $i = 4$  is the toe.



**Figure 1.** (a) A triaxial force sensor using a Cr–N thin film, and (b) schematic diagram of the structure of the sensor and layout of the Cr–N thin film.



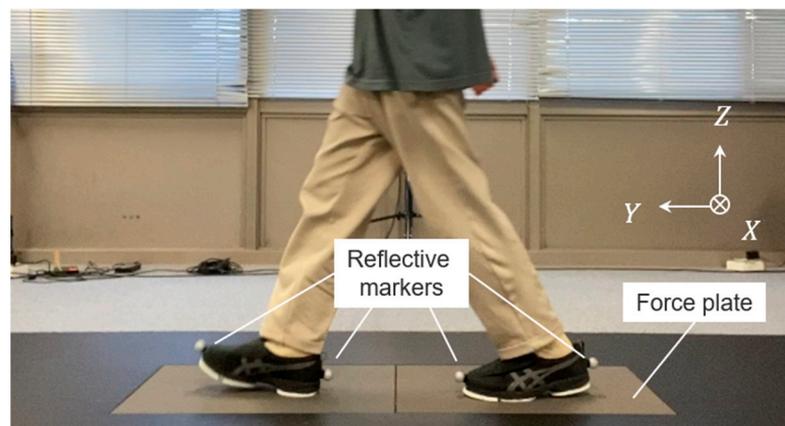
**Figure 2.** Sole sensor system. MT1 and MT5 represent the first and fifth metatarsal heads, respectively. (a) Location of the four triaxial force sensors, and (b) side view of the sole sensor system.

## 2.2. Participants

The study included six healthy young males and two healthy young females. The mean  $\pm$  SD age, height, and body mass of the participants were  $22.0 \pm 1.7$  years,  $1.69 \pm 0.048$  m, and  $58.0 \pm 5.4$  kg, respectively. The experimental protocol of this study was approved in advance by the Ethics Committee for Human Subjects Research, Graduate School of Engineering, Tohoku University (20A-5), and informed consent was obtained in writing from each participant after providing them with an explanation of the experimental methods and precautions in advance.

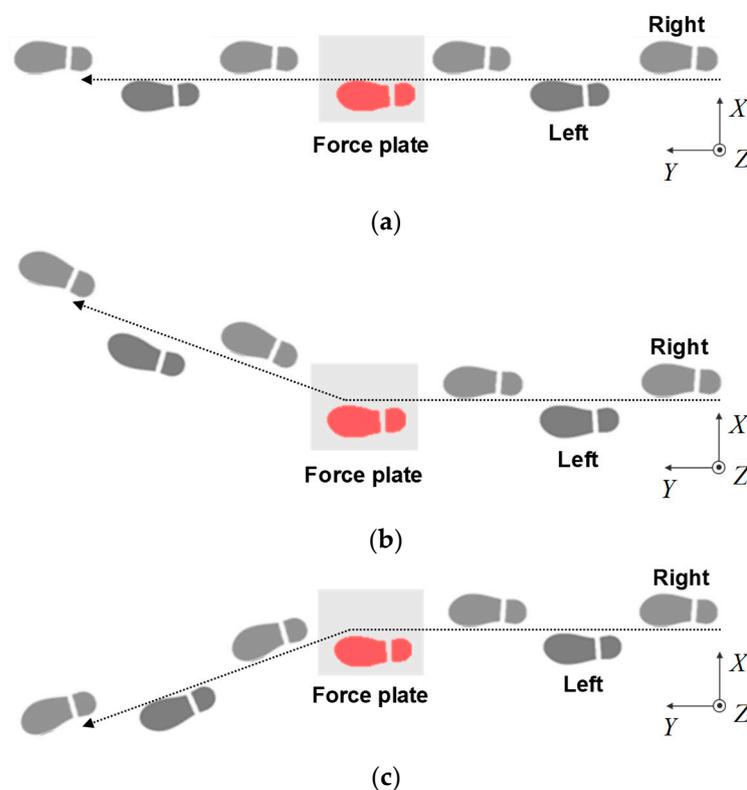
## 2.3. Experimental Procedure

Participants were instructed to walk on a 5 m long walkway with two force plates (FP4060-08, Bertec, Columbus, OH, USA; each was  $0.6 \text{ m} \times 0.4 \text{ m}$  in size) embedded in the center (Figure 3). Infrared reflective markers were attached to the toes and heels of the sole sensor system, and the position coordinates of the infrared reflective markers were measured using a three-dimensional motion analysis system (Optitrack, Acuity Inc., Reston, VA, USA). The sampling frequency for the position data of each marker in the three-dimensional motion analysis system was 200 Hz, and the sampling frequency of the GRF data ( $F_X$ ,  $F_Y$ , and  $F_Z$ ) on the force plate was 1000 Hz. The  $X$ ,  $Y$ , and  $Z$  coordinates in the walking experiment system were defined as shown in Figure 3, and the three-directional reaction forces measured by the force plate were denoted as  $F_X$ ,  $F_Y$ , and  $F_Z$ , respectively.



**Figure 3.** Experimental setup for gait trials, and the X, Y, and Z coordinates in the walking experiment system.

The participants were instructed to walk on a walkway from a stationary standing position with a self-selected stride and walking speed, and to step on the first force plate with the left foot. As shown in Figure 4, the participants were instructed to walk in a straight line and then turn on the first force plate. The participants performed two types of turns, a side-step turn (Figure 4b) and a cross-step turn (Figure 4c), as well as straight walking (Figure 4a), for a total of three types of walking styles. The turning angle was approximately 20 degrees for each turning trial. Participants performed multiple practice trials for each gait movement and performed 10 gait experiments for each gait movement.



**Figure 4.** Schematic of footprints for each type of gait trial. (a) Straight walking; (b) Side-step turning; (c) Cross-step turning.

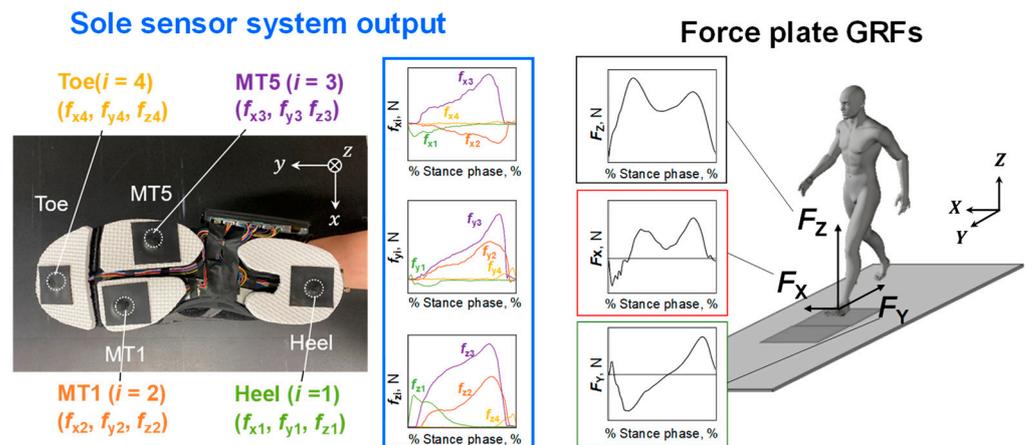
## 2.4. Data Analysis

### 2.4.1. Data Preprocessing

The time series data of  $f_{xi}$ ,  $f_{yi}$ , and  $f_{zi}$  ( $i = 1-4$ ), and  $F_x$ ,  $F_y$ , and  $F_z$  during the stance phase on the first force plate in each gait trial were used for the analysis (Figure 5). Matlab ver. 9.11 (Mathworks, Natick, MA, USA) was used for subsequent analyses. The time series data were smoothed by applying a fourth-order Butterworth low-pass filter with a cutoff frequency of 50 Hz. For the total value of  $f_{zi}$  ( $i = 1-4$ ), i.e.,  $\sum_{i=1}^4 f_{zi}$ , and the time series data of  $F_z$ , the threshold value was set at 15 N for the sole sensor system and 50 N for the force plate system [32,33] to determine the stance phase. The angle  $\varphi$  of the  $y$ -axis of the sensor to the  $Y$ -axis of the force plate coordinate axis was obtained from the reflective markers attached to the toes and heels, and the coordinate transformation for the GRF components of the force plate was performed as follows:

$$\begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix} = \begin{bmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} F_X \\ F_Y \\ F_Z \end{bmatrix} \quad (1)$$

where  $F_x$ ,  $F_y$ , and  $F_z$  are the GRF components obtained from the force plate in the coordinate system ( $x$ - $y$ - $z$ ) for the shoe sole sensor system. Thereafter, the time series data of both the shoe sole sensor system ( $f_{xi}$ ,  $f_{yi}$ , and  $f_{zi}$ ) and the force plate ( $F_x$ ,  $F_y$ , and  $F_z$ ) were reconstructed into 101 data sets (one for each 1%) by resampling, with the heel contact at 0% and the toe-off at 100%.



**Figure 5.** Forces obtained by the shoe sole sensor system and ground reaction forces (GRFs) obtained by a force plate. Twelve force sensor outputs ( $f_{xi}$ ,  $f_{yi}$ ,  $f_{zi}$  [ $i = 1-4$ ]) were used to estimate a model to predict the GRF in each direction.

### 2.4.2. Machine Learning Models

The test data  $T$  and training data  $D$  are expressed by the following equations:

$$T = [F_{\text{test}} | f_{\text{test}}] \quad (2)$$

$$F_{\text{test}} = [F_{x\text{test}} F_{y\text{test}} F_{z\text{test}}] \quad (3)$$

$$f_{\text{test}} = [f_{x1\text{test}} f_{x2\text{test}} f_{x3\text{test}} f_{x4\text{test}} f_{y1\text{test}} f_{y2\text{test}} f_{y3\text{test}} f_{y4\text{test}} f_{z1\text{test}} f_{z2\text{test}} f_{z3\text{test}} f_{z4\text{test}}] \quad (4)$$

$$D = [F | f] \quad (5)$$

$$F = [F_x F_y F_z] \quad (6)$$

$$\mathbf{f} = [f_{x1} f_{x2} f_{x3} f_{x4} f_{y1} f_{y2} f_{y3} f_{y4} f_{z1} f_{z2} f_{z3} f_{z4}] \quad (7)$$

$$[\mathbf{F}_{x\text{test}} \mathbf{F}_{y\text{test}} \mathbf{F}_{z\text{test}}] = \begin{bmatrix} F_{x\text{test}}^1 & F_{y\text{test}}^1 & F_{z\text{test}}^1 \\ \vdots & \vdots & \vdots \\ F_{x\text{test}}^j & F_{y\text{test}}^j & F_{z\text{test}}^j \\ \vdots & \vdots & \vdots \\ F_{x\text{test}}^{n_{\text{test}}} & F_{y\text{test}}^{n_{\text{test}}} & F_{z\text{test}}^{n_{\text{test}}} \end{bmatrix} \quad (8)$$

$$\mathbf{F}_x = \begin{bmatrix} F_x^1 & F_y^1 & F_z^1 \\ \vdots & \vdots & \vdots \\ F_x^j & F_y^j & F_z^j \\ \vdots & \vdots & \vdots \\ F_x^n & F_y^n & F_z^n \end{bmatrix} \quad (9)$$

$$[f_{xi\text{test}} f_{yi\text{test}} f_{zi\text{test}}] = \begin{bmatrix} f_{xi\text{test}}^1 & f_{yi\text{test}}^1 & f_{zi\text{test}}^1 \\ \vdots & \vdots & \vdots \\ f_{xi\text{test}}^j & f_{yi\text{test}}^j & f_{zi\text{test}}^j \\ \vdots & \vdots & \vdots \\ f_{xi\text{test}}^{n_{\text{test}}} & f_{yi\text{test}}^{n_{\text{test}}} & f_{zi\text{test}}^{n_{\text{test}}} \end{bmatrix} \quad (10)$$

$$[f_{xi} f_{yi} f_{zi}] = \begin{bmatrix} f_{xi}^1 & f_{yi}^1 & f_{zi}^1 \\ \vdots & \vdots & \vdots \\ f_{xi}^j & f_{yi}^j & f_{zi}^j \\ \vdots & \vdots & \vdots \\ f_{xi}^n & f_{yi}^n & f_{zi}^n \end{bmatrix} \quad (11)$$

where  $\mathbf{F}_{x\text{test}}$ ,  $\mathbf{F}_{y\text{test}}$ , and  $\mathbf{F}_{z\text{test}}$  are the  $x$ -,  $y$ -, and  $z$ -directional components of  $\mathbf{F}_{\text{test}}$ , respectively;  $f_{xi\text{test}}$ ,  $f_{yi\text{test}}$ , and  $f_{zi\text{test}}$  are the  $x$ -,  $y$ -, and  $z$ -directional components of  $\mathbf{f}_{\text{test}}$  at each sensor position ( $i = 1-4$ ), respectively;  $f_{xi}$ ,  $f_{yi}$ , and  $f_{zi}$  are the  $x$ -,  $y$ -, and  $z$ -directional components of  $\mathbf{f}$  at each sensor position ( $i = 1-4$ ), respectively; and  $n$  and  $n_{\text{test}}$  are the numbers of samples of the training data and test data, respectively.

Multiple linear regression (MLR) analysis was conducted using the GRFs obtained from the force plate as dependent variables and the triaxial forces obtained from the shoe sole sensor system as independent variables. The regression model is expressed by the following equation:

$$\begin{bmatrix} \hat{\mathbf{F}}_x \\ \hat{\mathbf{F}}_y \\ \hat{\mathbf{F}}_z \end{bmatrix} = \mathbf{f}_{\text{test}} \begin{bmatrix} \mathbf{k}_x \\ \mathbf{k}_y \\ \mathbf{k}_z \end{bmatrix} + \begin{bmatrix} \mathbf{b}_x \\ \mathbf{b}_y \\ \mathbf{b}_z \end{bmatrix} \quad (12)$$

where  $\hat{\mathbf{F}}_x$ ,  $\hat{\mathbf{F}}_y$ , and  $\hat{\mathbf{F}}_z$  are the vectors of estimated GRFs ( $n_{\text{test}} \times 1$ ) for a single step used in the test data;  $\mathbf{k}_x$ ,  $\mathbf{k}_y$ , and  $\mathbf{k}_z$  are the vectors of partial regression coefficients ( $k \times 1$ );  $\mathbf{b}_x$ ,  $\mathbf{b}_y$ , and  $\mathbf{b}_z$  are vectors consisting of constant terms ( $n_{\text{test}} \times 1$ ); and  $k$  is the number of independent variables. Partial regression coefficients were obtained using the least-squares method, and the independent variables used in the model were determined using the stepwise method.

Predictions were also made using Gaussian process regression (GPR), a nonlinear regression widely used in time series analysis [30]. GPR provides the probability distribution of the objective variable and outputs the prediction uncertainty as standard

deviation [28–31]. For example, in the  $x$  direction, the probability distribution of the predicted force plate GRF  $p(\hat{F}_x)$  is expressed in the following equation:

$$p(\hat{F}_x) = \mathcal{N}(F_{x\text{test}}^* + \mathbf{H}\boldsymbol{\beta}, \sigma^2 \mathbf{I}) \quad (13)$$

where  $F_{x\text{test}}^*$  is the latent variable in GPR,  $\mathbf{H}$  is a basis function vector consisting of 1,  $\boldsymbol{\beta}$  is the vector of the coefficients of  $\mathbf{H}$ ,  $\sigma$  is the noise standard deviation, and  $\mathbf{I}$  is the unit matrix. The forecasting procedure is as follows.

First, calculate  $\hat{\boldsymbol{\beta}}$  from the initial values of  $\boldsymbol{\theta}$  and  $\sigma^2$  computed from the basis matrix  $\mathbf{H}$  and the training data  $\mathbf{D}$ , using the following equation:

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}, \sigma^2) = \left[ \mathbf{H}^T [\mathbf{K} + \sigma^2 \mathbf{I}]^{-1} \mathbf{H} \right]^{-1} \mathbf{H}^T [\mathbf{K} + \sigma^2 \mathbf{I}]^{-1} \mathbf{F}_x \quad (14)$$

Then, estimate  $\boldsymbol{\theta}$  and  $\sigma^2$  that maximize the log-likelihood function expressed in the following equation:

$$\log p(\mathbf{F}_x) = -\frac{1}{2} (\mathbf{F}_x - \mathbf{H}\hat{\boldsymbol{\beta}})^T [\mathbf{K} + \sigma^2 \mathbf{I}]^{-1} (\mathbf{F}_x - \mathbf{H}\hat{\boldsymbol{\beta}}) - \frac{n}{2} \log 2\pi - \frac{1}{2} \log |\mathbf{K} + \sigma^2 \mathbf{I}| \quad (15)$$

Using the estimated  $\boldsymbol{\beta}$ ,  $\boldsymbol{\theta}$ , and  $\sigma^2$ , find the probability distribution of the latent variable  $F_{x\text{test}}^*$  in the GPR and calculate the probability distribution of the predicted values using the following formula:

$$\boldsymbol{\mu} = \mathbf{k}_* [\mathbf{K} + \sigma^2 \mathbf{I}]^{-1} \mathbf{F}_x - \mathbf{H}\boldsymbol{\beta} \quad (16)$$

$$\boldsymbol{\Sigma} = \mathbf{k}_{**} - \mathbf{k}_* [\mathbf{K} + \sigma^2 \mathbf{I}]^{-1} \mathbf{k}_*^T \quad (17)$$

$$p(\mathbf{F}_{x\text{test}}^* | \mathbf{f}_{\text{test}}, \mathbf{D}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (18)$$

where  $\mathbf{K}$ ,  $\mathbf{k}_*$ , and  $\mathbf{k}_{**}$  are kernel matrices and  $k(\mathbf{f}^i, \mathbf{f}^j)$  is the kernel function.

$$\mathbf{K} = \begin{bmatrix} k(\mathbf{f}^1, \mathbf{f}^1) & k(\mathbf{f}^1, \mathbf{f}^2) & \dots & k(\mathbf{f}^1, \mathbf{f}^n) \\ k(\mathbf{f}^2, \mathbf{f}^1) & k(\mathbf{f}^2, \mathbf{f}^2) & \dots & k(\mathbf{f}^2, \mathbf{f}^n) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{f}^n, \mathbf{f}^1) & k(\mathbf{f}^n, \mathbf{f}^2) & \dots & k(\mathbf{f}^n, \mathbf{f}^n) \end{bmatrix} \quad (19)$$

$$\mathbf{k}_* = \begin{bmatrix} k(\mathbf{f}^1, \mathbf{f}^1) & k(\mathbf{f}^1, \mathbf{f}^2) & \dots & k(\mathbf{f}^1, \mathbf{f}^{n_{\text{test}}}) \\ k(\mathbf{f}^2, \mathbf{f}^1) & k(\mathbf{f}^2, \mathbf{f}^2) & \dots & k(\mathbf{f}^2, \mathbf{f}^{n_{\text{test}}}) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{f}^{n_{\text{test}}}, \mathbf{f}^1) & k(\mathbf{f}^{n_{\text{test}}}, \mathbf{f}^2) & \dots & k(\mathbf{f}^{n_{\text{test}}}, \mathbf{f}^{n_{\text{test}}}) \end{bmatrix} \quad (20)$$

$$\mathbf{k}_{**} = \begin{bmatrix} k(\mathbf{f}^1, \mathbf{f}^1) & k(\mathbf{f}^1, \mathbf{f}^2) & \dots & k(\mathbf{f}^1, \mathbf{f}^{n_{\text{test}}}) \\ k(\mathbf{f}^2, \mathbf{f}^1) & k(\mathbf{f}^2, \mathbf{f}^2) & \dots & k(\mathbf{f}^2, \mathbf{f}^{n_{\text{test}}}) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{f}^{n_{\text{test}}}, \mathbf{f}^1) & k(\mathbf{f}^{n_{\text{test}}}, \mathbf{f}^2) & \dots & k(\mathbf{f}^{n_{\text{test}}}, \mathbf{f}^{n_{\text{test}}}) \end{bmatrix} \quad (21)$$

The squared exponential kernel was expressed as follows:

$$k(\mathbf{f}^i, \mathbf{f}^j) = \sigma_f^2 \exp \left[ -\frac{1}{2} \frac{(\mathbf{f}^i - \mathbf{f}^j)^T (\mathbf{f}^i - \mathbf{f}^j)}{\sigma_l^2} \right] \quad (22)$$

where  $\sigma_f$  and  $\sigma_l$  are the standard deviations of the training data of  $f$  and the characteristic length scale, respectively.  $\sigma_f$  and  $\sigma_l$  need to be greater than 0 and can be enforced by the unconstrained parametrization vector  $\theta = (\log \sigma_f, \log \sigma_l)^T$ . MLR and GRP models were trained using the Matlab ver. 9.11 (Mathworks, Natick, MA, USA) Statistics and Machine Learning Toolbox.

Prediction models for each movement and all movements were created by machine learning using all subject data. Therefore, 8 persons  $\times$  10 steps = 80 cases were used for each movement model, and 8 persons  $\times$  10 steps  $\times$  3 movements = 240 cases were used for all movement models. Leave-out-one cross-validation was used. That is, a regression model was created using one step of the measured data as test data  $T$  and the remaining data as training data  $D$ . By changing the trials used as test data and training data, the prediction accuracy of all trials used to create the model was determined, and the average of the predictions was calculated.

The mean absolute error (MAE), the %RMSE, and the degree-of-freedom-adjusted coefficient of determination  $R^2$  were used as evaluation indices for the prediction accuracy of the three-directional GRFs. The %RMSE was obtained by normalizing the RMSE by the range between the maximum and minimum values of the measured GRF values. For example, the MAE, %RMSE, and the  $R^2$  in the  $x$  direction are expressed by the following equations:

$$\text{MAE} = \frac{1}{n} \sum_{j=1}^n \left| F_{x\text{test}}^j - \hat{F}_{x\text{test}}^j \right| \quad (23)$$

$$\% \text{RMSE} = \sqrt{\frac{\sum_{j=1}^n (F_{x\text{test}}^j - \hat{F}_{x\text{test}}^j)^2}{n}} \frac{100}{F_{x\text{max}} - F_{x\text{min}}} \quad (24)$$

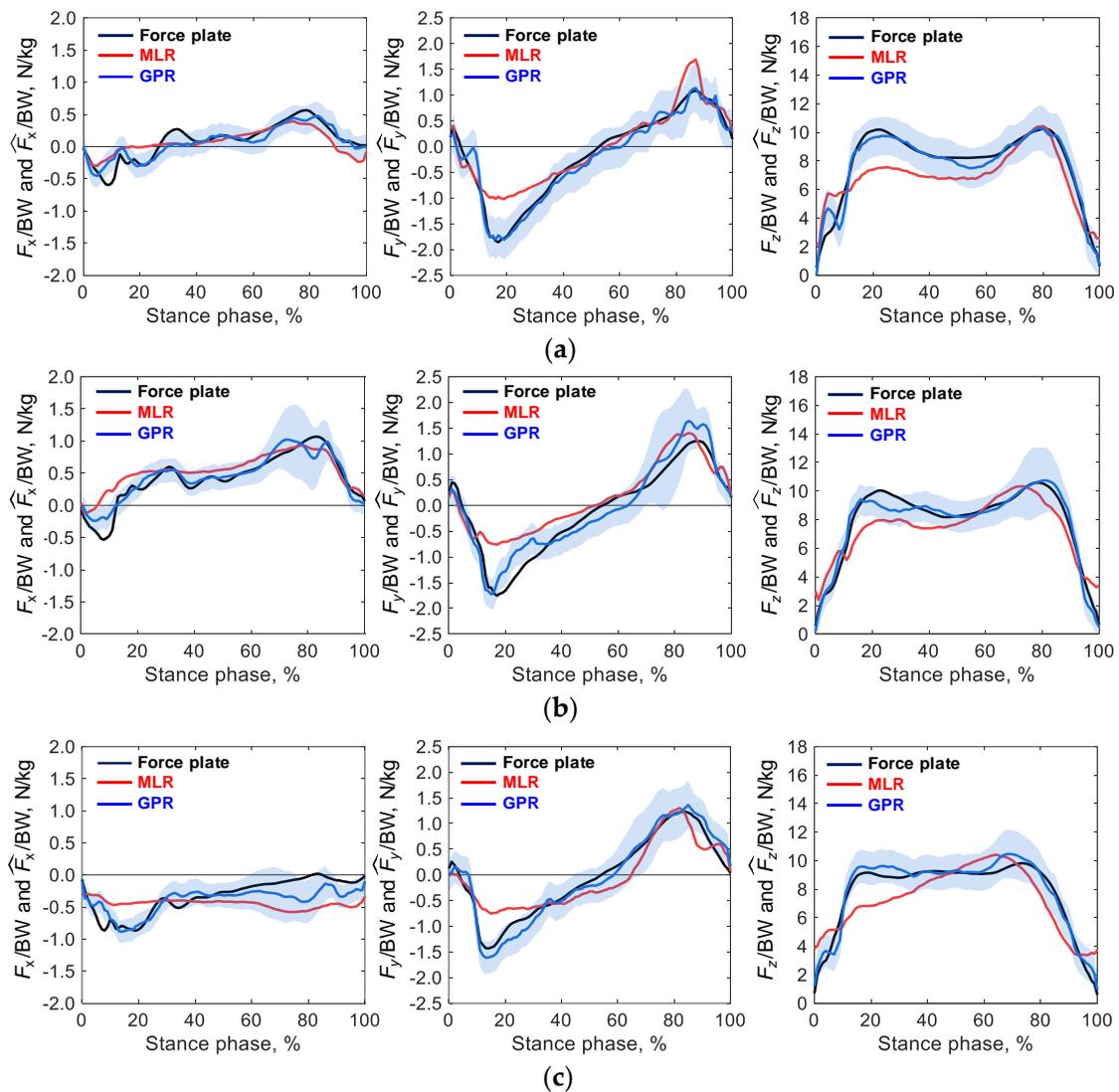
$$R^2 = 1 - \frac{\sum_{j=1}^n (F_{x\text{test}}^j - \hat{F}_{x\text{test}}^j)^2}{\sum_{j=1}^n (F_{x\text{test}}^j - \bar{F}_{x\text{test}})^2} \frac{n-1}{n-k-1} \quad (25)$$

where  $F_{x\text{test}}^j$  and  $\hat{F}_{x\text{test}}^j$  are the measured GRF used for the test data and the predicted GRF in the  $x$  direction ( $1 \leq j \leq n$ ), respectively, and  $\bar{F}_{x\text{test}}$  is the mean value of the measured GRF used for the test data. Note that the MAE was calculated using the  $F_{x\text{test}}^j$  and  $\hat{F}_{x\text{test}}^j$  divided by the participant's body mass.

### 3. Results

#### 3.1. Estimation of GRFs Using the Data of Each Movement

Figure 6 shows examples of time-series changes in the predicted and measured GRF values in the test trials for each movement, using the regression model trained on the data of each movement. The figure shows that in the  $y$  and  $z$  directions, the GPR model accurately predicted GRFs measured with the force plate throughout the stance phase for each movement. Conversely, the prediction using multiple regression analysis tended to be less accurate in the first half of the stance phase. Estimation accuracy was lower in cross-step turning (Figure 6c) than in straight walking (Figure 6a) and side-step turning (Figure 6b).



**Figure 6.** Examples of time-series changes in the predicted and measured ground reaction force (GRF) values in the test trials for each movement, using the regression model trained on the data of each movement for (a) straight walking, (b) side-step turning, and (c) cross-step turning. The horizontal axis indicates the normalized period with 0% for heel ground contact and 100% for toe-off, and the vertical axis indicates the measured GRFs and predicted GRFs divided by the participant's body mass. The solid black line shows the GRF values measured with the force plate, the solid red line shows the predictions by multiple linear regression (MLR), the solid blue line shows the predictions by Gaussian process regression (GPR), and the light blue shaded area shows the 95% confidence interval of the prediction by GRP.

Table 1 presents the mean  $\pm$  standard deviation of the prediction error, MAE, %RMSE, and  $R^2$  for each movement using the regression model trained on the data of all participants. The table shows that for both straight walking and side-step turning, the prediction of the GRF component in any direction had a %RMSE of less than 15%, regardless of the regression model. The GPR model had higher estimation accuracy than the MLR model, with a %RMSE of less than 15% in any direction of each type of gait, and the successful prediction was within the error range acceptable for clinical application. For straight walking and side-step turning, the GPR model showed good estimation accuracy, with a %RMSE of less than 10% and  $R^2$  of greater than 0.7, regardless of the direction. %RMSE and  $R^2$  was lower in the  $x$  direction than in the  $y$  and  $z$  directions; however, the MAE was lower in the  $x$  direction than in the  $y$  and  $z$  directions. In the case of cross-step turning, the

%RMSE exceeded 15% for the prediction of the GRF component in the  $x$  direction using the MLR model.

**Table 1.** Results of MAE, %RMSE, and  $R^2$  for each regression model trained on the data of each movement.

Type of Gait	Regression Model	$x$			$y$			$z$		
		MAE	%RMSE	$R^2$	MAE	%RMSE	$R^2$	MAE	%RMSE	$R^2$
Straight walking	MLR	0.067 ± 0.044	12.8 ± 4.3	0.551 ± 0.217	0.121 ± 0.104	9.5 ± 3.3	0.819 ± 0.071	0.617 ± 0.432	10.8 ± 4.5	0.682 ± 0.161
	GPR	0.047 ± 0.034	9.8 ± 2.7	0.706 ± 0.215	0.111 ± 0.090	6.2 ± 1.4	0.916 ± 0.066	0.237 ± 0.209	4.9 ± 2.2	0.917 ± 0.067
Side-step turning	MLR	0.102 ± 0.063	12.1 ± 2.6	0.673 ± 0.199	0.196 ± 0.160	11.3 ± 3.1	0.743 ± 0.193	0.460 ± 0.345	11.1 ± 4.4	0.674 ± 0.165
	GPR	0.072 ± 0.059	9.7 ± 2.2	0.755 ± 0.251	0.129 ± 0.187	7.5 ± 1.7	0.840 ± 0.458	0.195 ± 0.171	4.9 ± 1.7	0.922 ± 0.067
Cross-step turning	MLR	0.106 ± 0.092	20.1 ± 5.5	0.281 ± 1.629	0.178 ± 0.129	11.1 ± 2.9	0.752 ± 0.124	0.517 ± 0.340	12.3 ± 4.9	0.606 ± 0.176
	GPR	0.062 ± 0.055	13.8 ± 2.2	0.378 ± 0.541	0.124 ± 0.114	7.6 ± 1.2	0.867 ± 0.156	0.177 ± 0.186	5.1 ± 2.0	0.911 ± 0.107

Abbreviations: MAE, mean absolute error; %RMSE, percentage root-mean-square error; MLR, multiple linear regression; GPR, Gaussian process regression.

### 3.2. Estimation of GRFs Using Data of All Movements

Table 2 shows the mean ± standard deviation of the MAE, %RMSE, and  $R^2$  for each regression model trained on the data of all movements. The GPR model achieved lower MAE and %RMSE than the MLR model. Furthermore, the GPR model achieved a %RMSE of less than 10% in the  $y$  and  $z$  directions. However, in the  $x$  direction, the error was greater than 15% for both regression models. Thus, as with the models trained on each movement, the estimation accuracy was higher with the GPR model than with the MLR model. In addition, comparing these data with the data in Table 1, it was noted that the estimation accuracy was higher with the models trained specifically on the data of each movement than with the models trained on the data of all movements.

**Table 2.** Results of MAE, %RMSE and  $R^2$  for each regression model trained on the data of all movements.

Regression Model	$x$			$y$			$z$		
	MAE	%RMSE	$R^2$	MAE	%RMSE	$R^2$	MAE	%RMSE	$R^2$
MLR	0.222 ± 0.221	24.4 ± 6.6	−0.943 ± 2.478	0.181 ± 0.143	11.3 ± 3.1	0.749 ± 0.178	0.523 ± 0.383	11.8 ± 4.7	0.630 ± 0.176
GPR	0.120 ± 0.115	17.5 ± 3.5	−0.054 ± 1.501	0.115 ± 0.102	6.9 ± 1.0	0.894 ± 0.122	0.198 ± 0.168	5.2 ± 2.2	0.913 ± 0.063

Abbreviations: MAE, mean absolute error; %RMSE, percentage root-mean-square error; MLR, multiple linear regression; GPR, Gaussian process regression.

## 4. Discussion

In the three-directional GRF prediction using our shoe sole sensor system and machine learning, the %RMSE between the predicted and measured values was less than 15% for both regression models in the  $y$  and  $z$  directions, and the prediction accuracy of the GRFs was considered sufficient for practical use [6]. The error of the prediction using the model for straight walking only was smaller than that of the conventional three-directional GRF prediction using pressure sensors [20], indicating the practical feasibility of GRF prediction using our shoe sole sensor system. Alternatively, in the  $x$  direction, the %RMSE exceeded 17% for the prediction using the model of all movements, indicating that the prediction using the model of all movements is not sufficiently accurate.

### 4.1. Difference in Prediction Accuracy by Direction

Our results showed that the %RMSE for GRFs tended to be larger in the  $x$  direction than in the  $y$  and  $z$  directions, regardless of the type of training data, motion, or regression model used. However, the MAE in the  $x$  direction for each movement model was lower than that in the  $y$  and  $z$  directions. The MAE in the  $x$  direction for all movement model was comparable with that in the  $y$  direction and lower than that in the  $z$  direction. This is

because the absolute GRF values were larger in the  $y$  and  $z$  directions than in the  $x$  direction, resulting in a smaller %RMSE. In other studies [19,21], the %RMSE was calculated using the participant's body mass to normalize the RMSE value, and in this case, the %RMSE tended to be smaller in the  $x$  direction than in the  $y$  and  $z$  directions.

In addition, in this study, the directions of side-step turning and cross-step turning were reversed in all motion models, so the data of motion with GRFs in the opposite directions in the  $x$  direction were mixed. There was only one sensor in the heel, so the  $x$ -directional GRF output from the heel to the midfoot was not sufficient to distinguish between side-step and cross-step turning. Therefore, when the regression model was used for each movement, the prediction accuracy in the  $x$  direction for both side-step and cross-step turning improved compared with that in the all movement model, as shown in Tables 1 and 2. Based on these findings, it is considered effective to use different regression models to distinguish between different walking motions in order to improve the prediction accuracy of  $x$ -directional GRFs during different walking motions.

#### 4.2. Difference in Prediction Accuracy by Movement

As shown in Table 1, the prediction accuracy in the  $x$  direction tended to be lower for cross-step turning than for straight walking and side-step turning, regardless of the regression model. Cross-step turning is a difficult movement in which the body's center of mass tends to deviate from the base of support area and the base of support is narrow, making it necessary to change direction while maintaining body balance during the movement [34,35]. We determined the between-participant and within-participant standard deviations of the GRFs divided by the participant's body mass throughout the stance phase in the  $x$  direction. The mean value of the between-participant standard deviation of the GRF during cross-step turning (0.221) was larger than that during straight walking (0.148) and side-step turning (0.151). The mean value of the within-participant standard deviation of the GRF during cross-step turning (0.183) was larger than that during straight walking (0.089) and side-step turning (0.121). Therefore, the variation of GRFs in the left-right direction during cross-step turning within-participant and between participants was larger than that during straight walking and side-step turning. The prediction accuracy for side-step turning was equivalent to that for straight walking, as shown in Table 1. This is thought to be due to the fact that the side-step turning has a larger base of support in the left-right direction [34,35], making it easier to maintain body balance in the left-right direction.

#### 4.3. Difference in Prediction Accuracy by Regression Models

As shown in Tables 1 and 2, the error was smaller in the GPR model than in the MLR model, as expected. GPR, which is a nonlinear regression, allows the construction of complex models dealing with nonlinear relationships between objective variables and explanatory variables that cannot be handled by multiple regression analysis, which is a linear regression. In this study, we did not use methods, such as neural networks [36], because of the small data set, but such machine learning methods are effective if the training data set is made larger by increasing the number of participants and the number of steps. However, the results of this study indicate that GPR can perform accurate learning even with relatively small data sets.

#### 4.4. Study Limitations

Some limitations of the current study should be considered. First, the number of participants was small, and the representation of genders was uneven. Thus, our results may not be generalizable to a broader population. In addition, studies with a large sample size will provide improved GRF prediction accuracy. Second, there were only three types of walking movements. It is considered necessary to create prediction models with a data set that considers various movements performed in daily life in addition to straight walking, side-step turning, and cross-step turning.

## 5. Conclusions

The regression model trained for each motion, except for MLR in cross-step turning, showed less than 15% prediction error (%RMSE) regardless of the direction. The three-directional GRFs during walking predicted by the shoe sole sensor system showed relatively good agreement with the GRFs measured with the force plate, indicating the practical applicability of the shoe sole sensor system for gait analysis. However, the prediction error in the  $x$  direction (%RMSE) for the whole motion model exceeded 15%. Compared with straight walking and side-step turning, cross-step turning tended to have a lower prediction accuracy in the  $x$  direction, regardless of the regression model.

The prediction of three-directional GRFs using our shoe sole sensor system and machine learning could be used in a wide range of practical applications, such as in the sports field and for gait analysis at rehabilitation facilities and in daily life, taking advantage of its portability and accuracy in prediction.

**Author Contributions:** Conceptualization, T.Y. and Y.T.; methodology, T.Y., Y.T. and Y.S.; software, Y.T.; validation, T.Y. and Y.T.; formal analysis, T.Y. and Y.T.; investigation, T.Y. and Y.T.; data curation, T.Y.; writing—original draft preparation, T.Y.; writing—review and editing, T.Y., Y.T. and Y.S.; supervision, T.Y.; project administration, T.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The experimental protocol for this study was approved in advance by the Ethics Committee for Human Subjects Research, Graduate School of Engineering, Tohoku University (20A-5).

**Informed Consent Statement:** Written informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data including graphs and table within this paper are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Laroche, D.P.; Cook, S.B.; Mackala, K. Strength asymmetry increases gait asymmetry and variability in older women. *Med. Sci. Sports Exerc.* **2012**, *44*, 2172–2181. [[CrossRef](#)]
2. Hsiao, H.; Gray, V.L.; Creath, R.A.; Binder-Macleod, S.A.; Rogers, M.W. Control of lateral weight transfer is associated with walking speed in individuals post-stroke. *J. Biomech.* **2017**, *60*, 72–78. [[CrossRef](#)]
3. Franz, J.R.; Maletis, M.; Kram, R. Real-time feedback enhances forward propulsion during walking in old adults. *Clin. Biomech.* **2014**, *29*, 68–74. [[CrossRef](#)] [[PubMed](#)]
4. Eltoukhy, M.; Kuenze, C.; Andersen, M.S.; Oh, J.; Signorile, J. Prediction of ground reaction forces for Parkinson's disease patients using a kinect-driven musculoskeletal gait analysis model. *Med. Eng. Phys.* **2017**, *50*, 75–82. [[CrossRef](#)] [[PubMed](#)]
5. Skals, S.; Rasmussen, K.P.; Bendtsen, K.M.; Yang, J.; Andersen, M.S. A musculoskeletal model driven by dual Microsoft Kinect sensor data. *Multibody Syst. Dyn.* **2017**, *41*, 297–316. [[CrossRef](#)]
6. Liu, M.M.; Herzog, W.; Savelberg, H.H.C.M. Dynamic muscle force predictions from EMG: An artificial neural network approach. *J. Electromyogr. Kinesiol.* **1999**, *9*, 391–400. [[CrossRef](#)]
7. Shahabpoor, E.; Pavic, A. Estimation of vertical walking ground reaction force in real-life environments using single IMU sensor. *J. Biomech.* **2018**, *79*, 181–190. [[CrossRef](#)] [[PubMed](#)]
8. Martínez-Pascual, D.; Catalán, J.M.; Blanco-Ivorra, A.; Sanchís, M.; Arán-Ais, F.; García-Aracil, N. Estimating vertical ground reaction forces during gait from lower limb kinematics and vertical acceleration using wearable inertial sensors. *Front. Bioeng. Biotechnol.* **2023**, *29*, 1199459. [[CrossRef](#)]
9. Patoz, A.; Lussiana, T.; Breine, B.; Gindre, C.; Malatesta, D. A Single Sacral-Mounted Inertial Measurement Unit to Estimate Peak Vertical Ground Reaction Force, Contact Time, and Flight Time in Running. *Sensors* **2022**, *22*, 784. [[CrossRef](#)]
10. Tan, T.; Chiasson, D.P.; Hu, H.; Shull, P.B. Influence of IMU position and orientation placement errors on ground reaction force estimation. *J. Biomech.* **2019**, *97*, 109416. [[CrossRef](#)]
11. Bamberg, S.J.M.; Benbasat, A.Y.; Scarborough, D.M.; Krebs, D.E.; Paradiso, J.A. Gait analysis using a shoe-integrated wireless sensor system. *IEEE Trans. Inf. Technol. Biomed.* **2008**, *12*, 413–423. [[CrossRef](#)]
12. Sazonov, E.S.; Fulk, G.; Hill, J.; Schutz, Y.; Browning, R. Monitoring of posture allocations and activities by a shoe-based wearable sensor. *IEEE Trans. Bio-Med. Eng.* **2011**, *58*, 983–990. [[CrossRef](#)]

13. Redd, C.B.; Bamberg, S.J.M. A wireless sensory feedback device for real-time gait feedback and training. *IEEE/ASME Trans. Mechatron.* **2012**, *17*, 425–433. [[CrossRef](#)]
14. Dyer, P.S.; Bamberg, S.J.M. Instrumented insole vs. force plate: A comparison of center of plantar pressure. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Boston, MA, USA, 30 August–3 September 2011; Volume 2011, pp. 6805–6809. [[CrossRef](#)]
15. Howell, A.M.; Kobayashi, T.; Chou, T.R.; Daly, W.; Orendurff, M.; Bamberg, S.J.M. A laboratory insole for analysis of sensor placement to determine ground reaction force and ankle moment in patients with stroke. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, San Diego, CA, USA, 28 August–1 September 2012; Volume 2012, pp. 6394–6397. [[CrossRef](#)]
16. Fong, D.T.P.; Chan, Y.Y.; Hong, Y.; Yung, P.S.H.; Fung, K.Y.; Chan, K.M. Estimating the complete ground reaction forces with pressure insoles in walking. *J. Biomech.* **2008**, *41*, 2597–2601. [[CrossRef](#)]
17. Howell, A.M.; Kobayashi, T.; Hayes, H.A.; Foreman, K.B.; Bamberg, S.J.M. Kinetic gait analysis using a low-cost insole. *IEEE Trans. Bio-Med. Eng.* **2013**, *60*, 3284–3290. [[CrossRef](#)] [[PubMed](#)]
18. Pu, F.; Yang, Y.; Fan, X.; Li, S.; Li, Y.; Li, D.; Fan, Y. Optimal estimation of total plantar force for monitoring gait in daily life activities with low-price insole system. *J. Mech. Med. Biol.* **2014**, *14*, 1450037. [[CrossRef](#)]
19. Rouhani, H.; Favre, J.; Crevoisier, X.; Aminian, K. Ambulatory assessment of 3D ground reaction force using plantar pressure distribution. *Gait Posture* **2010**, *32*, 311–316. [[CrossRef](#)] [[PubMed](#)]
20. Jacobs, D.A.; Ferris, D.P. Estimation of ground reaction forces and ankle moment with multiple, low-cost sensors. *J. NeuroEngineering Rehabil.* **2015**, *12*, 90. [[CrossRef](#)]
21. Sim, T.; Kwon, H.; Oh, S.E.; Joo, S.B.; Choi, A.; Heo, H.M.; Kim, K.; Mun, J.H. Predicting complete ground reaction forces and moments during gait with insole plantar pressure information using a wavelet neural network. *J. Biomech. Eng.* **2015**, *137*, 091001. [[CrossRef](#)]
22. Schepers, H.M.; Van Asseldonk, E.H.F.; Buurke, J.H.; Veltink, P.H. Ambulatory estimation of center of mass displacement during walking. *IEEE Trans. Bio-Med. Eng.* **2009**, *56*, 1189–1195. [[CrossRef](#)]
23. Adachi, W. Development of walking analysis system using by motion sensor with mobile force plate. *Trans. JSME* **2012**, *78*, 1607–1616. [[CrossRef](#)]
24. Moriyasu, K.; Nishiwaki, T.; Yamaguchi, T.; Hokkirigawa, K. New technique of three directional ground reaction force distributions. *Footwear Sci.* **2010**, *2*, 57–64. [[CrossRef](#)]
25. Yamaguchi, T. Distribution of the local required coefficient of friction in the shoe–floor contact area during straight walking: Pilot study. *Biotribology* **2019**, *19*, 100101. [[CrossRef](#)]
26. Niwa, E.; Shirakawa, K.; Shingyochi, S.; Xiong, S.; Nakahara, K.; Ito, T.; Sasaki, Y. Load vector sensors using strain sensitive cr-n thin films and their application. *IEEJ Trans. Sens. Micromach.* **2014**, *134*, 132–139. [[CrossRef](#)]
27. Niwa, E.; Sasaki, Y. Cr-N strain sensitive thin films and their pressure sensor applications. *IEEJ Trans. Sens. Micromach.* **2014**, *134*, 385–391. [[CrossRef](#)]
28. Rasmussen, C.E. Gaussian processes in machine learning. *Lect. Notes Comput. Sci.* **2004**, *3176*, 63–71. [[CrossRef](#)]
29. Eguchi, R.; Takahashi, M. Insole-based estimation of vertical ground reaction force using one-step learning with probabilistic regression and data augmentation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 1217–1225. [[CrossRef](#)]
30. Wen, Y.; Li, B.; Chen, D.; Zhu, T. Reliability and validity analysis of personality assessment model based on gait video. *Front. Behav. Neurosci.* **2022**, *16*, 901568. [[CrossRef](#)]
31. Raja, M.N.A.; Abdoun, T.; El-Sekelly, W. Smart prediction of liquefaction-induced lateral spreading. *J. Rock Mech. Geotech. Eng.* **2023**; *in press*. [[CrossRef](#)]
32. Burnfield, J.M.; Power, C.M. The role of center of mass kinematics in predicting peak utilized coefficient of friction during walking. *J. Forensic Sci.* **2007**, *52*, 1328–1333. [[CrossRef](#)]
33. Yamaguchi, T.; Yano, M.; Onodera, H.; Hokkirigawa, K. Kinematics of center of mass and center of pressure predict friction requirement at shoe-floor interface during walking. *Gait Posture* **2013**, *38*, 209–214. [[CrossRef](#)] [[PubMed](#)]
34. Hase, K.; Stein, R.B. Turning strategies during human walking. *J. Neurophysiol.* **1999**, *81*, 2914–2922. [[CrossRef](#)]
35. Taylor, M.J.D.; Dabnichhi, P.; Strike, S.C. A three-dimensional biomechanical comparison between turning strategies during the stance phase of walking. *Hum. Mov. Sci.* **2005**, *24*, 558–573. [[CrossRef](#)]
36. Honert, E.C.; Hoitz, F.; Blades, S.; Nigg, S.R.; Nigg, B.M. Estimating Running Ground Reaction Forces from Plantar Pressure during Graded Running. *Sensors* **2022**, *22*, 3338. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.