

Article

Feature Optimization for Gait Phase Estimation with a Genetic Algorithm and Bayesian Optimization

Wonseok Choi [†] , Wonseok Yang [†], Jaeyoung Na, Giuk Lee  and Woochul Nam ^{*} 

Departments of Mechanical Engineering, Chung-Ang University, Seoul 06974, Korea; tnqr2765@gmail.com (W.C.); yws1108@gmail.com (W.Y.); tommyna@naver.com (J.N.); giuklee@cau.ac.kr (G.L.)

^{*} Correspondence: wcnam@cau.ac.kr; Tel.: +82-2-820-5270

[†] These authors contributed equally to this work as first authors.

Abstract: For gait phase estimation, time-series data of lower-limb motion can be segmented according to time windows. Time-domain features can then be calculated from the signal enclosed in a time window. A set of time-domain features is used for gait phase estimation. In this approach, the components of the feature set and the length of the time window are influential parameters for gait phase estimation. However, optimal parameter values, which determine a feature set and its values, can vary across subjects. Previously, these parameters were determined empirically, which led to a degraded estimation performance. To address this problem, this paper proposes a new feature extraction approach. Specifically, the components of the feature set are selected using a binary genetic algorithm, and the length of the time window is determined through Bayesian optimization. In this approach, the two optimization techniques are integrated to conduct a dual optimization task. The proposed method is validated using data from five walking and five running motions. For walking, the proposed approach reduced the gait phase estimation error from 1.284% to 0.910%, while for running, the error decreased from 1.997% to 1.484%.

Keywords: gait phase; feature optimization; time-domain feature; time window; genetic algorithm; Bayesian optimization



Citation: Choi, W.; Yang, W.; Na, J.; Lee, G.; Nam, W. Feature Optimization for Gait Phase Estimation with a Genetic Algorithm and Bayesian Optimization. *Appl. Sci.* **2021**, *11*, 8940. <https://doi.org/10.3390/app11198940>

Academic Editor: Manuel Armada

Received: 21 July 2021

Accepted: 23 September 2021

Published: 25 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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

Walking and running are essential motions in daily life [1]. Gait motions can be analyzed by gait phases because gait motion is a cyclic movement composed of several gait phases [2]. Therefore, gait phase identification is widely used to evaluate and diagnose gait. Gait phases have been investigated using foot pressure sensors [3] and inertial measurement units (IMUs) [4]. Accordingly, gait phase analysis has been used in several fields, such as rehabilitation [5], diagnosis [6,7], and assistive robot design [8,9]. The performance of these applications relies on the gait phase estimation accuracy.

Gait phase estimation has been conducted using two different approaches: the discrete and continuous gait phase approaches. In the discrete gait phase approach, the gait event is recognized as one of the discrete gait phases, such as loading response, mid-stance, and terminal stance [10]. These discrete phases are differentiated based on the kinematic characteristics of lower limbs. In the continuous gait phase approach, the gait phase is a motion status that is defined as a value between 0% and 100% of the gait cycle [8]. More specifically, the heel strike of the left (or right) foot matches with 0% and 100%. Other gait events within one cycle correspond to values between 0% and 100%; the gait phase values linearly increase from 0% to 100% over time during one cycle. In practical applications, the discrete gait phase approach has drawbacks when compared with the continuous gait phase approach. First, assigning gait states to discrete gait phases is challenging, especially in transition states between two distinct phases [11]. This gait phase misidentification may provide unreliable gait information, which results in inadequate controls. Furthermore, controls that are determined by the discrete phase

approach lead to unsmooth controls at transition states because different phases require different controls [12]. Therefore, the continuous gait phase approach is suitable for accurate and precise controls in various applications.

Gait phase estimation algorithms can utilize either original time-series signals or statistical features obtained from time-series signals. To obtain these statistical features, the time-series data from two consecutive instants, or times, are segmented. Then, the statistical features (e.g., the mean and standard deviation (SD) of the segmented time-series data) can be calculated. These statistical features are referred to as time-domain (TD) features, and the segmentation length is referred to as the length of the time window (LTW). Estimation models based on the original time-series signals [13–17] can capture the temporal patterns of such signals. However, complicated neural networks are required to train the temporal patterns from the original time-series data [18]. Moreover, this end-to-end learning approach requires large datasets to achieve high accuracy. TD features can address these issues because the number of features is very small compared to that of the original time-series data.

If TD features are used for gait estimation, the temporal characteristics of time-series signals can be represented by a relatively small number of features [19,20]. The TD features have been proven to obtain high performance with a low computational burden [21]; thus, they have been widely used in gait phase estimation. Kang et al. [22] used a sliding window to obtain several TD features from IMU signals (e.g., mean, variance, minimum, and maximum). Subsequently, the continuous gait phase was estimated via a multi-layer perceptron (MLP), which was trained using the extracted TD features. Farah et al. [23] investigated various gait phase classifiers that used a TD feature set calculated from the angle of the knee and the angular velocity and acceleration of the thigh. Caramia et al. [24] introduced several TD feature sets derived from eight IMU datasets. Their proposed TD feature sets were utilized to classify the gait patterns of healthy subjects and of patients suffering from Parkinson's disease. Mazilu et al. [25] calculated TD features from accelerometer signals and conducted a principal component analysis to detect the freezing of gait of patients suffering from Parkinson's disease. In addition to the aforementioned studies, several gait phase estimations with TD features have been adopted and performed satisfactorily [26–31].

Because gait patterns vary across subjects, it is not appropriate to use the same TD feature set and an identical LTW for different subjects. For example, the SD value of the thigh angle is a critical feature in the estimation of the gait phase of a specific subject, whereas this feature may be redundant for another subject [32]. Furthermore, the TD feature values depend on the LTW. Even if gait patterns are similar across subjects, the feature values may be different across the same subjects owing to variations in stride time [33,34]. Therefore, the selection of an optimal TD feature set and assignment of the LTW are crucial in gait phase estimation.

This paper proposes an optimization approach to effectively determine the TD feature set and LTW for MLP-based gait estimation. The TD features were determined using a binary genetic algorithm (BGA). A BGA is an effective feature selection technique that is widely used in many applications [35–37]. This algorithm not only selects important features but also minimizes the dimension of the feature set. After the TD features are selected, the LTW can be optimized via Bayesian optimization (BO), a hyperparameter-tuning method that can rapidly converge with a small number of iterations [38]. The LTW was optimized by minimizing the gait phase error, which is estimated from the MLP. The BGA/BO combination is realized through iterative processes, where the BGA chooses the candidate TD feature set, and BO searches for an optimal LTW value for the candidate feature set. This method repeats this calculation until the BGA finds an optimal TD feature set.

The remainder of this paper is organized as follows. Section 2 describes the thigh IMU signal used in this study for gait phase estimation. Section 3 describes the MLP-based gait phase estimation model used in this study. Subsequently, the proposed method for

the feature set and hyperparameter determination is explained. Section 4 presents the optimized results for each of the walking and running cases evaluated. Finally, the paper is concluded in Section 5.

2. Data Description

This study used the IMU data of gait motions from five healthy male subjects (28 ± 3 years; 181.3 ± 8.0 cm; 78.7 ± 9.8 kg, mean $\pm$ SD); these IMU data were obtained from a previously conducted experiment in a previous work [39]. In that experiment, two IMU sensors (MTi-3 AHRS, Xsens Technologies, the Netherlands) were attached to the anterior part of each thigh to collect the data, as shown in Figure 1. The subjects walked (1.5 m/s) and ran (2.5 m/s) on an instrumented treadmill with mounted force plates (Bertec, Columbus, OH, USA). The IMU sensors recorded the thigh angles and angular velocities at a 1 kHz sampling rate. Details regarding the size of the signal used in this study, including the number of strides and the total experimental time for each subject, are provided in Table 1. Seventy percent of the time-series data were used for training the MLP, the remaining 30% of the data were used for testing.

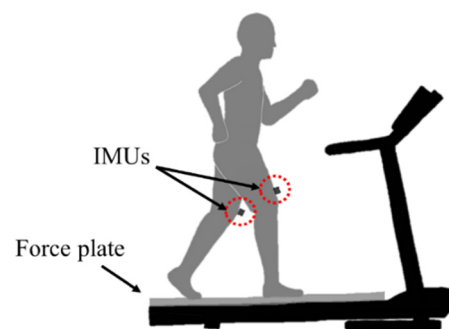


Figure 1. Experimental setup for the acquisition of the gait motion data.

Table 1. Details of the IMU signal used in this study.

Subject	Motion	Number of Strides	Time Lengths (s)
1	Walking	12	12.210
	Running	29	22.885
2	Walking	18	18.480
	Running	16	11.305
3	Walking	15	13.905
	Running	20	13.815
4	Walking	29	29.240
	Running	24	18.810
5	Walking	22	21.975
	Running	12	9.240

The thigh motion IMU data of the five subjects are shown in Figure 2. The data show variations in gait patterns across subjects. Because of these variations, if the same feature set and LTW are used for every subject, the performance of the gait estimation algorithm may downgrade.

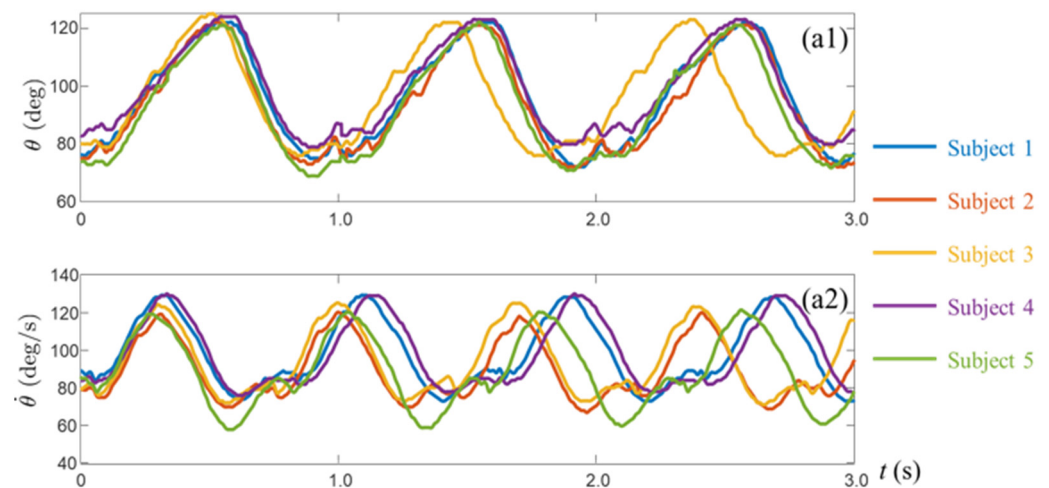


Figure 2. (a1) Angle and (a2) angular velocity data of left thigh motion for the five subjects involved in this study.

3. Gait Phase Estimation Model

3.1. MLP for Gait Phase Estimation

This section describes the gait estimation procedure using TD features. First, the TD features are calculated from the original time-series data in the current time window, as shown in Figure 3. Notably, the LTW is different for each TD feature set. Notably, the feature values were not normalized because a previous study [22] showed that an accurate estimation can be obtained without normalization.

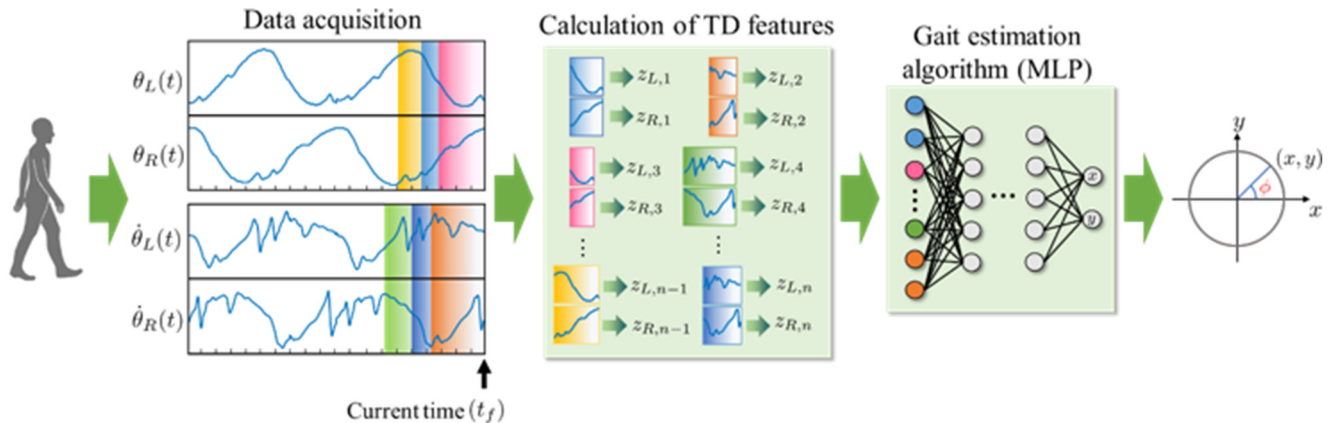


Figure 3. Gait phase estimation procedure. $z_{L,i}$ and $z_{R,i}$ are the i -th TD features of the left and right lower limb, respectively.

Next, the calculated TD features are fed into an MLP consisting of two hidden layers. Because the optimal selection of TD features is one of the main objectives of this study, the number of input nodes can vary depending on the number of selected TD features. The two hidden layers were constructed with 15 and 20 nodes, respectively. In this study, $\tanh$ activation function was applied to every hidden layer. The details on the MLP architecture are provided in Tables S4 and S5 in Supplementary Materials. The two output nodes represent continuously varying variables x and y , which can be converted to the gait phase p_g as:

$$\begin{aligned} x &= \cos \phi, \\ y &= \sin \phi, \\ \text{where } \phi &= \frac{p_g}{100} \times 2\pi \end{aligned} \quad (1)$$

If the error in the gait phase ϕ is directly used in the regression cost function, the result would be unsatisfactory. Because the gait phase discontinuously changed from 100% to

0% at the end of each gait cycle, the error is much larger near 0% (and 100%) than at other points. To avoid this unequally distributed error, a conversion technique (1) was adapted from a previous study [22]. The error is very large around the end of the gait cycle. For example, if the ground truth is 1%, the MLP prediction lags behind by 2%. Then, the MLP result would be 99%, and the absolute error is 98%. Thus, this unrealistic error must be addressed using (1).

This issue can be addressed by converting the gait phase into continuously varying variables x and y , as shown in (1). After the MLP calculates x and y , the current gait phase is obtained using the $\arctan2$ function. A loss function L is defined using the mean absolute error as follows:

$$L = \frac{1}{2n} \sum_{i=1}^n (|x_i - \hat{x}_i| + |y_i - \hat{y}_i|), \quad (2)$$

where n represents the number of training data points. The Adam optimizer was applied with a learning rate of 0.001 and 5000 iteration epochs, which were fixed throughout the study. The estimation algorithm was trained on a workstation with a single graphics processing unit (CPU: Intel Xeon CPU E5-2620 v4, 2.10 GHz, and 16 cores; GPU: NVIDIA GeForce RTX 2080 Ti). TensorFlow was used as the deep learning framework.

3.2. Optimization Algorithms

The main objectives of this study are (1) the selection of the best feature set and (2) obtaining the optimal LTWs. A new optimizing procedure is proposed to achieve these objectives simultaneously, as shown in Figure 4. The BGA selects the feature sets, and BO determines the optimal LTWs. Using this structure, both the TD features and LTWs can be effectively optimized in a single procedure.

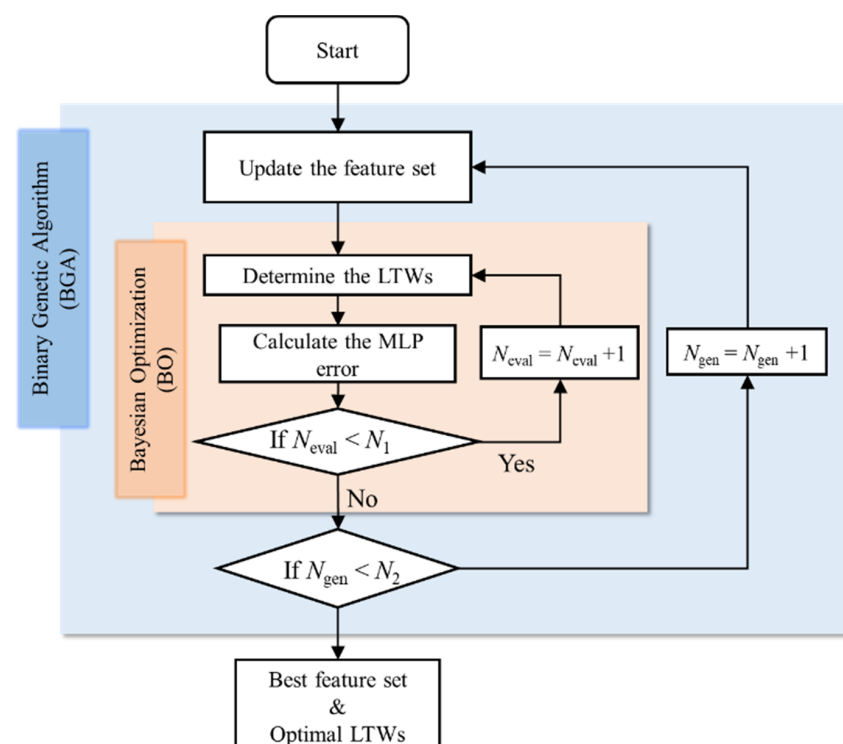


Figure 4. Feature optimization procedure.

To conduct this optimization, the estimation performance should be quantified. Therefore, the estimation performance was evaluated using the estimation error as follows:

$$\text{Error} = \frac{\text{RMSE}}{2} \times 100 (\%) \quad (3)$$

where

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2}{2n}}.$$

Here, x_i and y_i are the MLP-predicted values, and $\hat{x}_i$ and $\hat{y}_i$ are the true values calculated from the true gait phase values. The root mean square error (RMSE) is divided by two because the range of x_i and y_i is between -1 and 1 . Seventy percent of the time-series data were used for training the MLP. Subsequently, the test error was computed using the remaining 30% of the data. The test error was used as the objective function to be minimized in the optimization procedure.

3.2.1. Binary Genetic Algorithm for Feature Selection

A BGA was used to select an optimal feature set for the target subject. The BGA generates different populations iteratively to create binary-type genes, each of which corresponds to one of the TD features. The BGA can be implemented using the following procedure. First, an initial population is generated. Then, the objective function value is calculated for each chromosome; the objective function is described in Section 3.2.2. The chromosome is a vector consisting of binary values, and each binary value implies the presence of a designated TD feature. Next, a new population is created by applying crossover and mutation to the chromosomes. Finally, the objective functions of the new population are evaluated. This procedure can be repeated until the generation number reaches a predefined value.

A feature pool for the BGA was constructed with 12 TD features: minimum, maximum, mean (μ), standard deviation (σ), initial (t_i), and final (t_f) values for the thigh angle and angular velocity. Specifically, the feature pool was constructed as $(\theta_{\min}, \theta_{\max}, \mu_{\theta}, \sigma_{\theta}, \theta_{(t_i)}, \theta_{(t_f)}, \dot{\theta}_{\min}, \dot{\theta}_{\max}, \mu_{\dot{\theta}}, \sigma_{\dot{\theta}}, \dot{\theta}_{(t_i)}, \dot{\theta}_{(t_f)})$. If zero is allocated to a gene, the corresponding feature is not used in the gait phase estimation. If unity is allocated to a gene, the corresponding feature is used in the estimation. All features are composed of the values for the left and right thighs. For example, if μ_{θ} is selected as a feature, then the mean values of both the left and right thigh angles (i.e., $\mu_{L,\theta}$ and $\mu_{R,\theta}$) are used in the gait phase estimation, as shown in Figure 5.

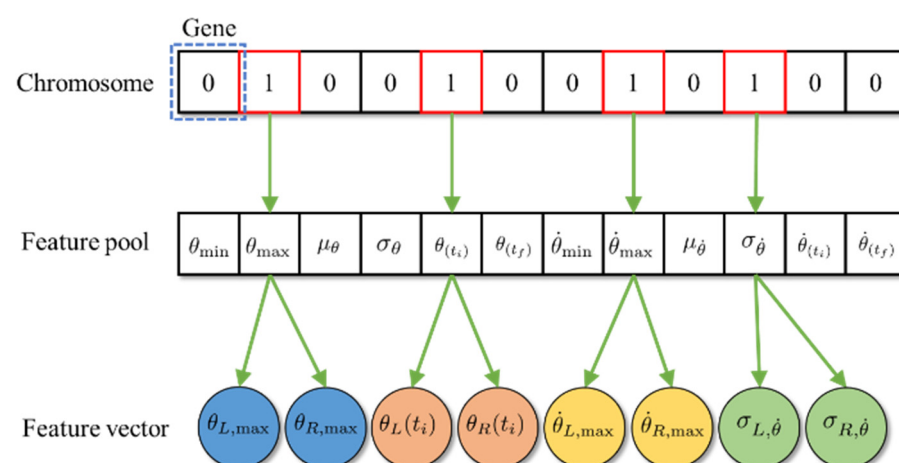


Figure 5. TD feature selection in the BGA. TD features in the feature pool are selected or rejected when the binary values in the chromosome are determined. For example, if the “1” is allocated in the second element of the chromosome, $\theta_{\max}$ is used as the feature. Then, $\theta_{\max}$ of both the right thigh and left thigh motion are used for gait phase estimation.

In this study, the population size for each generation was fixed at 15, the maximum number of generations at 20, the crossover rate at 0.5, and the mutation rate at 0.05. Although Grefenstette et al. [40] suggested using 50 initial populations, this study selected

15 initial populations considering the small dimension of the feature pool. Consequently, the other parameters were adjusted based on [40]. Additionally, roulette-wheel crossover and bit-flip mutation were applied to the crossover and mutation methods, respectively. The roulette wheel is a widely used stochastic selection method in which the probability of an individual being selected is proportional to its fitness value [41]. This strategy enables the delivery of valuable individuals to the next generation with a high probability. The bit-flip mutation flips the binary value of an individual to generate a mutated chromosome [42]. Finally, one elite chromosome for each generation was preserved and passed to the next generation.

3.2.2. Bayesian Optimization for LTW

In this study, BO was used as the optimizer for the LTWs of the TD features, and the loss values of the test datasets, obtained by (2), were the values to be minimized. The BO of these loss values was realized using a surrogate model and an acquisition function. A Gaussian process was used as the surrogate model, and an expected improvement function was used as the acquisition function. The parameters of the BO algorithm implemented in this study are summarized in Table 2.

Table 2. BO parameters.

Bayesian Optimization Parameter	Value
Number of initial points	5
Number of iterations	20
Search space of parameters	(100, 800)
Surrogate model	Gaussian process
Kernel function	Matern 5/2
Acquisition function	Expected Improvement
Explore-exploit trade-off parameter	0.01

After the BGA assigns a feature set to the BO, the BO determines the LTW value of each TD feature. Subsequently, the MLP is trained with the features calculated using the LTW, as shown in Figure 6. Next, the MLP returns the performance of the test dataset. Then, the BO selects new LTW values based on the previous results. The procedure described above is repeated until a predefined number of iterations are conducted. Notably, in this study, the BO was conducted with five random initial points for every feature set to obtain reliable results. Through this procedure, the BO is able to obtain the optimal gait estimation performance for the assigned feature set, and the corresponding optimal performance is returned to the BGA.

There are two special cases in which the BO needs to be operated differently. First, if every gene in a chromosome is zero, the MLP cannot estimate the gait phase. Second, if the selected TD feature set consists only of the final angle $\theta(t_f)$ and angular velocity $\dot{\theta}(t_f)$ value, the LTW determination is unnecessary because these final values are not affected by the LTW. Thus, BO does not iteratively optimize the LTWs in these cases.

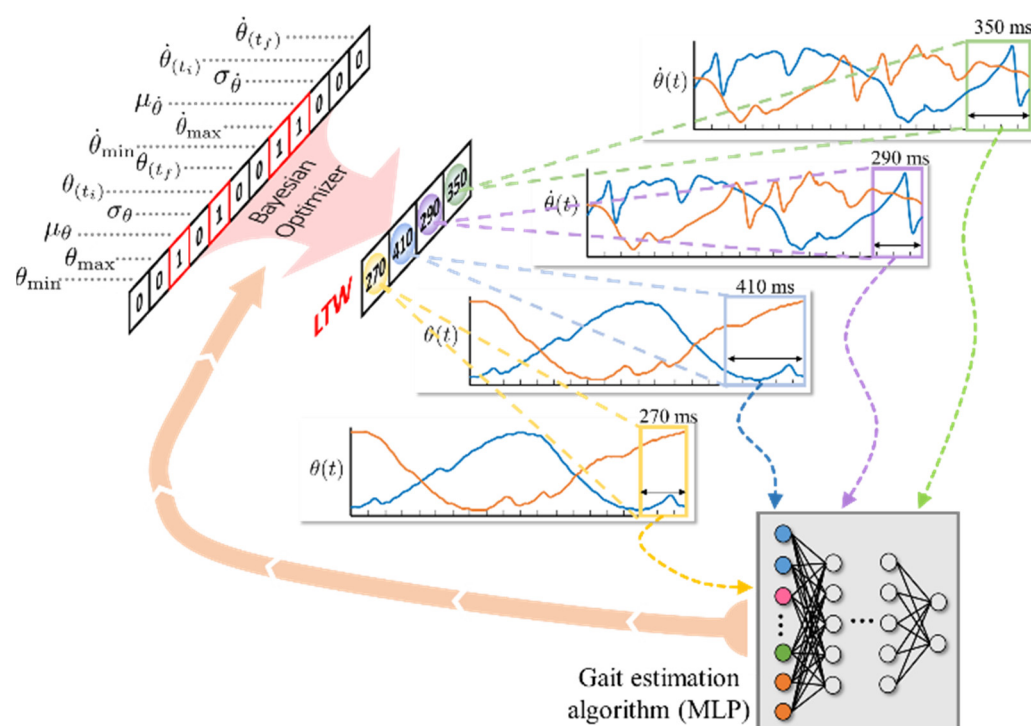


Figure 6. Example of independent LTW selection with BO. In this example, $\mu_{\dot{\theta}}$, $\theta_{(t_i)}$, $\theta_{(t_f)}$, $\dot{\theta}_{\max}$, and $\mu_{\dot{\theta}}$ are chosen as the feature set, and BO optimizes the LTW as 270, 410, 290, and 350 ms, respectively.

4. Results

The proposed algorithm was evaluated using the walking and running motion data of five subjects. Because this study aims to optimize the gait estimation algorithm for each specific subject, the error of the proposed optimization was compared to that of a heuristic model.

4.1. Estimation Error of the Heuristic Model

A feature set and LTW are empirically determined in the heuristic model. Five feature sets were selected, as shown in Table 3. Feature set 1 is a set comprising 12 TD features; feature sets 2 and 3 consist of TD features extracted from the angle and angular velocity, respectively; feature sets 4 and 5 are randomly selected from the possible TD feature set. When training the MLP with feature sets 1–5, the LTW value was set to 300 ms, as suggested in a previous study [22].

Table 3. Feature sets for comparison.

Feature set 1	$\theta_{\min}, \theta_{\max}, \mu_{\theta}, \sigma_{\theta}, \theta_{(t_i)}, \theta_{(t_f)}, \dot{\theta}_{\max}, \dot{\theta}_{\min}, \mu_{\dot{\theta}}, \sigma_{\dot{\theta}}, \dot{\theta}_{(t_i)}, \dot{\theta}_{(t_f)}$
Feature set 2	$\theta_{\min}, \theta_{\max}, \mu_{\theta}, \sigma_{\theta}, \theta_{(t_i)}, \theta_{(t_f)}$
Feature set 3	$\dot{\theta}_{\max}, \dot{\theta}_{\min}, \mu_{\dot{\theta}}, \sigma_{\dot{\theta}}, \dot{\theta}_{(t_i)}, \dot{\theta}_{(t_f)}$
Feature set 4	$\theta_{\max}, \theta_{\min}, \sigma_{\theta}, \dot{\theta}_{\max}, \dot{\theta}_{\min}, \sigma_{\dot{\theta}}$
Feature set 5	$\mu_{\theta}, \theta_{(t_i)}, \theta_{(t_f)}, \mu_{\dot{\theta}}, \dot{\theta}_{(t_i)}, \dot{\theta}_{(t_f)}$

The MLP was trained using a heuristically determined feature set and LTW. Seventy percent of the time-series data were used for training. The test error was computed using the remaining 30% of the data. The MLP was trained using five different initial values for weights and biases. The mean and standard deviations of the five errors were obtained and are presented in Tables 4 and 5. For walking, the error ranges from 1.1% to 1.6% in most cases, as presented in Table 4. The error is relatively large when feature set 3 is used,

suggesting that angle information is essential for gait phase estimation. For running, the error is approximately 2%, which is relatively larger than that for walking, as shown in Table 5.

Table 4. Mean estimation error of heuristic feature sets for walking (No optimization). The numbers in parenthesis represent the standard deviation.

Walking	Subject ID					Average
	1	2	3	4	5	
Feature set 1	1.488% (0.083)	1.206% (0.078)	1.301% (0.078)	1.408% (0.099)	1.228% (0.035)	1.326
Feature set 2	1.225% (0.059)	1.388% (0.079)	1.151% (0.078)	1.385% (0.016)	1.456% (0.046)	1.321
Feature set 3	2.150% (0.344)	2.000% (0.198)	2.047% (0.283)	2.075% (0.463)	2.256% (0.164)	2.106
Feature set 4	1.624% (0.067)	1.524% (0.105)	1.553% (0.078)	1.583% (0.072)	1.681% (0.145)	1.593
Feature set 5	1.264% (0.071)	1.258% (0.078)	1.297% (0.055)	1.458% (0.045)	1.238% (0.039)	1.303

Table 5. Mean estimation error of heuristic feature sets for running (no optimization). The numbers in parenthesis represent the standard deviation.

Running	Subject ID					Average
	1	2	3	4	5	
Feature set 1	2.250% (0.113)	1.919% (0.102)	2.678% (0.082)	1.687% (0.117)	2.113% (0.106)	2.129
Feature set 2	2.254% (0.060)	2.085% (0.052)	2.545% (0.033)	1.503% (0.093)	1.493% (0.096)	1.976
Feature set 3	2.539% (0.124)	2.554% (0.079)	2.904% (0.104)	1.556% (0.115)	2.064% (0.278)	2.323
Feature set 4	2.317% (0.187)	2.207% (0.384)	2.895% (0.099)	1.860% (0.114)	2.301% (0.128)	2.316
Feature set 5	2.388% (0.058)	2.055% (0.080)	2.617% (0.116)	1.697% (0.163)	1.515% (0.108)	2.054

To quantitatively compare the performance of the heuristic model and the proposed optimizer, one of the five heuristic feature sets was selected and used for comparison. For walking, feature set 1 provides the smallest or second smallest error. Therefore, this set was used for the comparison. Similarly, feature set 2 was chosen as the representative heuristic set for running.

4.2. Error Reduction Due to Optimization

During the optimization procedure, the error value decreased and converged for all subjects, as shown in Figure 7. The final optimization results for walking and running are presented in Tables 6 and 7, respectively. The shaded boxes indicate that the corresponding feature was not included in the best feature set. While the initial angle $\theta(t_i)$ and angular velocity $\dot{\theta}(t_i)$ depend on the LTW, the final angle $\theta(t_f)$ and angular velocity $\dot{\theta}(t_f)$ are independent of the LTW. Thus, no LTW values are provided for $\theta(t_f)$ and $\dot{\theta}(t_f)$.

To assess the error reduction for walking, the error for feature set 1 is also provided in Table 6 (see the numbers in parentheses in the last row). For walking, when the error is averaged over all subjects, the average error is 0.910% after optimization. This subject-average error is 1.284% when feature set 1 is used. For running, the error for feature set 2 is provided in Table 7 (see the numbers in parentheses in the last row). The average error is 1.997% for feature set 2. The error decreased to 1.484% when using the proposed approach.

Additional case study is provided in Tables S1 and S2 in Supplementary Materials, which suggests the importance of feature selection on the estimation accuracy. It is worth noting that the optimal LTW is considerably different between some subjects. This difference is caused by the local minima of error over LTW. Details are provided in Figure S2 in Supplementary Materials.

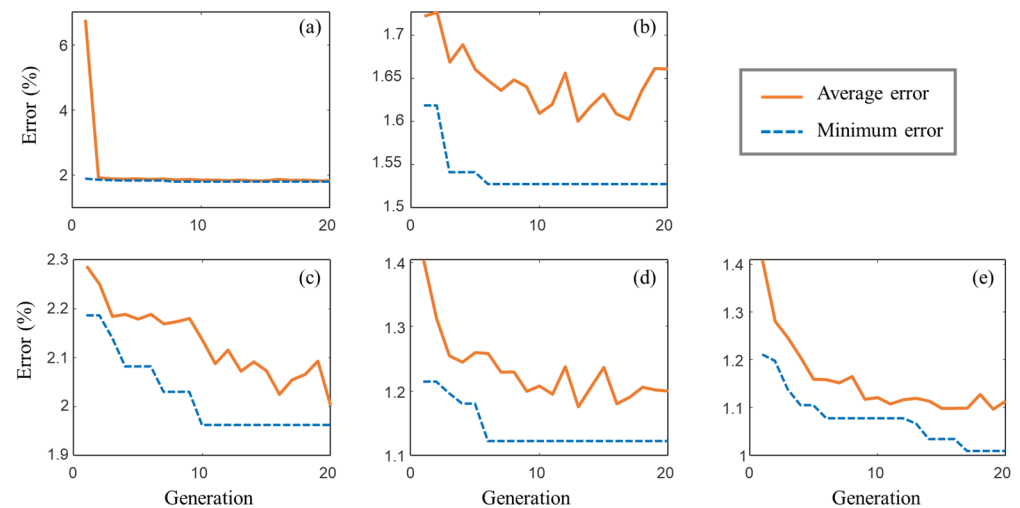


Figure 7. Optimization results over generations for walking. (a–e) shows the average and minimum errors for subjects 1–5, respectively.

Table 6. Optimized feature set, LTW, and error for walking.

Walking	Subject ID				
	1	2	3	4	5
$\theta_{\min}$	345 ms	790 ms		183 ms	178 ms
$\theta_{\max}$	209 ms			200 ms	
μ_{θ}	638 ms	301 ms	325 ms	761 ms	557 ms
σ_{θ}	367 ms	395 ms		465 ms	663 ms
$\theta(t_i)$		571 ms	477 ms		227 ms
$\theta(t_f)$	-		-	-	-
$\dot{\theta}_{\min}$		385 ms		390 ms	102 ms
$\dot{\theta}_{\max}$		373 ms	342 ms	285 ms	
$\mu_{\dot{\theta}}$		738 ms		642 ms	186 ms
$\sigma_{\dot{\theta}}$			478 ms		
$\dot{\theta}(t_i)$		253 ms	704 ms	548 ms	
$\dot{\theta}(t_f)$	-			-	
Error after optimization	0.920%	0.780%	0.933%	1.045%	0.874%
(Error of feature set 1)	(1.523%)	(1.206%)	(1.199%)	(1.321%)	(1.169%)

A gray shadow represents a feature that is not selected in the best feature set. A hyphen indicates that this feature is selected, but no LTW value is required to calculate the corresponding feature value.

Table 7. Optimized feature set, LTW, and error for running.

Running	Subject ID				
	1	2	3	4	5
$\theta_{\min}$		574 ms		636 ms	383 ms
$\theta_{\max}$	184 ms				363 ms
μ_{θ}	782 ms				
σ_{θ}		312 ms	727 ms	157 ms	799 ms
$\theta(t_i)$		336 ms	135 ms	678 ms	633 ms
$\theta(t_f)$	-	-	-	-	
$\dot{\theta}_{\min}$	314 ms	365 ms			
$\dot{\theta}_{\max}$				109 ms	
$\mu_{\dot{\theta}}$	458 ms	318 ms		225 ms	408 ms
$\sigma_{\dot{\theta}}$	208 ms				432 ms
$\dot{\theta}(t_i)$	186 ms	688 ms	501 ms	370 ms	
$\dot{\theta}(t_f)$	-				-
Error after optimization (Error of feature set 2)	1.798% (2.225%)	1.527% (2.108%)	1.962% (2.560%)	1.122% (1.593%)	1.009% (1.497%)

A gray shadow represents a feature that is not selected in the best feature set. A hyphen indicates that this feature is selected, but no LTW value is required to calculate the corresponding feature value.

5. Discussion

In this study, the best feature set and its LTW were determined for each subject with BGA and BO. Some TD features were excluded from the best feature set, suggesting that some features were less useful than others. For example, the optimizer selects only six features among the 12 candidate features for subjects 1 and 3 for walking. For running, only four features were selected for subject 3. These small-sized feature sets show higher accuracy than other feature sets, which shows that optimization can also reduce the number of features. A smaller number of features reduces the computation time required for feature extraction. Table 8 lists the computation time for feature extraction. The computation cost for σ_{θ} is considerably larger than that for the other features. Thus, a feature set that does not contain σ_{θ} would be much faster than other feature sets with σ_{θ} . Although the computational speed was not considered for optimization in this study, it could be included in future optimization studies.

Table 8. Comparison of the computation time required for feature extraction with an LTW of 300 ms. The ratio represents the computation time normalized by the computation time required to extract $\theta(t_i)$.

	$\theta(t_i)$	$\theta(t_f)$	$\theta_{\min}$	$\theta_{\max}$	μ_{θ}	σ_{θ}
Ratio	1	1	22.15	21.88	28.24	34.14

This calculation was conducted using a workstation with a single graphics processing unit (CPU: Intel Xeon CPU E5-2620 v4, 2.10 GHz, and 16 cores; GPU: NVIDIA GeForce RTX 2080 Ti).

The best feature sets and optimal LTW values varied considerably across subjects, as shown in Tables 6 and 7. This finding implies that personal variations in features and LTWs are considerable; thus, personalized optimization is necessary for gait estimation. Furthermore, the best feature sets for walking and running are different, which indicates that the effects of features on the estimation are also affected by gait speed. Because the gait patterns of walking and running motions may differ considerably, the gait phase needs to be estimated using different TD features.

This study proposes a mathematical procedure to extract the TD features and LTWs. Previously, most studies selected a TD feature set without a systematic approach. For example, Kang et al. [22] used a single TD feature set for each subject. Farah et al. [23]

changed the feature sets manually and compared the results. The proposed method is especially effective when the number of candidate features is very large. A grid search is a traditional feature selection approach in which all possible cases are verified to find an optimal set. However, this grid search requires a considerable amount of time if the number of features is large. Considering the time required for optimization, the proposed method is much more efficient than the grid search approach for large feature pools.

6. Conclusions

The gait phase estimation performance strongly depends on features. However, parameter determination for feature extraction is complicated. For gait estimation, TD feature set selection and LTW determination require different techniques. Previous studies determined the features by trial and error, which does not guarantee optimal features. To address this problem, this study proposes an optimizing procedure composed of a BGA and BO. This approach can effectively reduce the estimation error by simultaneously considering the feature set and LTWs. This study also revealed that the best feature set and optimal LTW vary across subjects and motion types (i.e., walking and running).

The proposed approach can be modified depending on its application. For example, if the gait phase is required for assistive robots, computation speed is the most important factor. In this system, the upper bound of the LTW can be lowered during the optimization process. Then, the time required to calculate the values of the TD features can be decreased. The maximum number of TD features can also be defined in order to reduce computation time. Penalty values can be allocated to individual features. Some features require a very short computation time, whereas other features require significant computational resources. For instance, $\theta(t_i)$ and $\theta(t_f)$ can be handled with very small computational resources. However, μ_θ and σ_θ require a relatively lengthy computation time, as shown in Table 8. Thus, the cost function of the optimizer can be modified to consider computation time. For example, a high cost can be applied if μ_θ and σ_θ are selected because their computation time is relatively long. This modification can provide different features that are optimized for computational speed.

Feature optimization for various walking (or running) speeds is required for practical use. In this study, optimization was conducted for gait motions where the speed remained constant over time. As shown in the optimization results, the optimal feature set and LTW changes according to walking (or running) speed. Thus, if the gait phase needs to be estimated at various speeds, the optimizer should consider the estimation errors over the target speed range. Alternatively, adaptive feature selection could be used as a scheme to handle speed variations. Walking (or running) speed can be classified into discrete levels. Then, the optimal feature set and LTW at each level is predefined. Subsequently, the estimation algorithm is able to handle speed variations by updating the feature set and LTW according to the walking (or running) speed.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/app11198940/s1>, Table S1: Effects of optimization on the test error for walking, Table S2: Effects of optimization on the test error for running, Table S3: Estimation result with different MLP nodes for walking, Table S4: Estimation result with different MLP nodes for running, Figure S1: BO result for the LTW of $\theta_{\min}$, Reference [22].

Author Contributions: Conceptualization, W.N.; methodology, W.C. and W.Y.; software, J.N.; validation, W.C. and W.Y.; formal analysis, W.C.; investigation, W.C.; resources, W.C.; data curation, W.Y.; writing—original draft preparation, W.C. and W.Y.; writing—review and editing, W.N. and G.L.; visualization, W.Y. and W.C.; supervision, W.C.; project administration, W.C. and G.L.; funding acquisition, G.L. and W.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1A4A3030268) and the Chung-Ang University Research Scholarship Grants in 2021.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: This study has used data measured in a previously published study.

Data Availability Statement: The data supporting the results of this study are available from the corresponding author upon reasonable request.

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

References

- Brandes, M.; Schomaker, R.; Möllenhoff, G.; Rosenbaum, D. Quantity versus quality of gait and quality of life in patients with osteoarthritis. *Gait Posture* **2008**, *28*, 74–79. [[CrossRef](#)] [[PubMed](#)]
- Rueterbories, J.; Spaich, E.G.; Larsen, B.; Andersen, O.K. Methods for gait event detection and analysis in ambulatory systems. *Med. Eng. Phys.* **2010**, *32*, 545–552. [[CrossRef](#)]
- 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](#)] [[PubMed](#)]
- Seel, T.; Raisch, J.; Schauer, T. IMU-Based Joint Angle Measurement for Gait Analysis. *Sensors* **2014**, *14*, 6891–6909. [[CrossRef](#)] [[PubMed](#)]
- Lünenburger, L.; Colombo, G.; Riener, R. Biofeedback for robotic gait rehabilitation. *J. Neuroeng. Rehabil.* **2007**, *4*, 1. [[CrossRef](#)] [[PubMed](#)]
- Simon, S.R. Quantification of human motion: Gait analysis—benefits and limitations to its application to clinical problems. *J. Biomech.* **2004**, *37*, 1869–1880. [[CrossRef](#)]
- Mariani, B.; Jimenez, M.C.; Vingerhoets, F.J.; Aminian, K. On-shoe wearable sensors for gait and turning assessment of patients with Parkinson's disease. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 155–158. [[CrossRef](#)]
- Aoyagi, D.; Ichinose, W.; Harkema, S.; Reinkensmeyer, D.; Bobrow, J. An Assistive Robotic Device That Can Synchronize to the Pelvic Motion During Human Gait Training. In Proceedings of the 9th International Conference on Rehabilitation Robotics, ICORR 2005, Chicago, IL, USA, 28 June–1 July 2005; pp. 565–568.
- Yan, T.; Parri, A.; Garate, V.R.; Cempini, M.; Ronsse, R.; Vitiello, N. An oscillator-based smooth real-time estimate of gait phase for wearable robotics. *Auton. Robot.* **2017**, *41*, 759–774. [[CrossRef](#)]
- Taborri, J.; Palermo, E.; Rossi, S.; Cappa, P. Gait Partitioning Methods: A Systematic Review. *Sensors* **2016**, *16*, 66. [[CrossRef](#)]
- Villarreal, D.J.; Poonawala, H.A.; Gregg, R.D. A Robust Parameterization of Human Gait Patterns Across Phase-Shifting Perturbations. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 265–278. [[CrossRef](#)]
- Xu, D.; Crea, S.; Vitiello, N.; Wang, Q. Capacitive Sensing-Based Continuous Gait Phase Estimation in Robotic Transfemoral Prostheses. In Proceedings of the 2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob), New York, NY, USA, 29 November–1 December 2020; pp. 298–303.
- Jung, J.-Y.; Heo, W.; Yang, H.; Park, H. A Neural Network-Based Gait Phase Classification Method Using Sensors Equipped on Lower Limb Exoskeleton Robots. *Sensors* **2015**, *15*, 27738–27759. [[CrossRef](#)] [[PubMed](#)]
- Sui, J.-D.; Chen, W.-H.; Shiang, T.-Y.; Chang, T.-S. Real-Time Wearable Gait Phase Segmentation for Running and Walking. In Proceedings of the 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Seville, Spain, 12–14 October 2020; pp. 1–5.
- Seo, K.; Park, Y.J.; Lee, J.; Hyung, S.; Lee, M.; Kim, J.; Choi, H.; Shim, Y. RNN-Based On-Line Continuous Gait Phase Estimation from Shank-Mounted IMUs to Control Ankle Exoskeletons. In Proceedings of the 2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR), Toronto, ON, Canada, 24–28 June 2019; Volume 2019, pp. 809–815.
- Ding, Z.; Yang, C.; Xing, K.; Ma, X.; Yang, K.; Guo, H.; Yi, C.; Jiang, F. The Real Time Gait Phase Detection Based on Long Short-Term Memory. In Proceedings of the 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), Guangzhou, China, 18–21 June 2018; pp. 33–38.
- Wang, S.; Cao, J.; Yu, P. Deep Learning for Spatio-Temporal Data Mining: A Survey. *IEEE Trans. Knowl. Data Eng.* **2020**. [[CrossRef](#)]
- Koutnik, J.; Greff, K.; Gomez, F.; Schmidhuber, J. A Clockwork RNN. In Proceedings of the International Conference on Machine Learning 2014, Beijing, China, 21–26 June 2014; pp. 1863–1871.
- Phinyomark, A.; Phukpattaranont, P.; Limsakul, C. Feature reduction and selection for EMG signal classification. *Expert Syst. Appl.* **2012**, *39*, 7420–7431. [[CrossRef](#)]
- Wahab, N.I.A.; Mohamed, A.; Hussain, A. Fast transient stability assessment of large power system using probabilistic neural network with feature reduction techniques. *Expert Syst. Appl.* **2011**, *38*, 11112–11119. [[CrossRef](#)]
- Boostani, R.; Moradi, M.H. Evaluation of the forearm EMG signal features for the control of a prosthetic hand. *Physiol. Meas.* **2003**, *24*, 309–319. [[CrossRef](#)]
- Kang, I.; Kunapuli, P.; Young, A.J. Real-time neural network-based gait phase estimation using a robotic hip exoskeleton. *IEEE Trans. Med. Robot. Bionics* **2019**, *2*, 28–37. [[CrossRef](#)]
- Farah, J.D.; Baddour, N.; Lemaire, E.D. Gait Phase Detection from Thigh Kinematics Using Machine Learning Techniques. In Proceedings of the 2017 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Rochester, MN, USA, 7–10 May 2017; pp. 263–268.

24. Caramia, C.; Torricelli, D.; Schmid, M.; Munoz-Gonzalez, A.; Gonzalez-Vargas, J.; Grandas, F.; Pons, J.L. IMU-Based Classification of Parkinson's Disease from Gait: A Sensitivity Analysis on Sensor Location and Feature Selection. *IEEE J. Biomed. Health Inform.* **2018**, *22*, 1765–1774. [[CrossRef](#)]
25. Mazilu, S.; Calatroni, A.; Gazit, E.; Roggen, D.; Hausdorff, J.M.; Tröster, G. Feature Learning for Detection and Prediction of Freezing of Gait in Parkinson's Disease. In Proceedings of the Programming Languages and Systems, Rome, Italy, 16–24 March 2013; pp. 144–158.
26. Meng, M.; She, Q.; Gao, Y.; Luo, Z. EMG Signals Based Gait Phases Recognition Using Hidden Markov Models. In Proceedings of the 2010 IEEE International Conference on Information and Automation, Harbin, China, 20–23 June 2010; pp. 852–856. [[CrossRef](#)]
27. Meng, M.; Luo, Z.; She, Q.; Ma, Y. Automatic recognition of gait mode from EMG signals of lower limb. In Proceedings of the 2010 The 2nd International Conference on Industrial Mechatronics and Automation, Wuhan, China, 30–31 May 2010; Volume 1, pp. 282–285.
28. Sejdic, E.; Lowry, K.A.; Bellanca, J.; Redfern, M.S.; Brach, J.S. A Comprehensive Assessment of Gait Accelerometry Signals in Time, Frequency and Time-Frequency Domains. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 603–612. [[CrossRef](#)] [[PubMed](#)]
29. Wang, L.; Sun, Y.; Li, Q.; Liu, T.; Yi, J. Two Shank-Mounted IMUs-Based Gait Analysis and Classification for Neurological Disease Patients. *IEEE Robot. Autom. Lett.* **2020**, *5*, 1970–1976. [[CrossRef](#)]
30. Hsu, W.-C.; Sugiarto, T.; Lin, Y.-J.; Yang, F.-C.; Lin, Z.-Y.; Sun, C.-T.; Hsu, C.-L.; Chou, K.-N. Multiple-Wearable-Sensor-Based Gait Classification and Analysis in Patients with Neurological Disorders. *Sensors* **2018**, *18*, 3397. [[CrossRef](#)] [[PubMed](#)]
31. Peng, F.; Peng, W.; Zhang, C.; Zhong, D. IoT Assisted Kernel Linear Discriminant Analysis Based Gait Phase Detection Algorithm for Walking With Cognitive Tasks. *IEEE Access* **2019**, *7*, 68240–68249. [[CrossRef](#)]
32. Pierrynowski, M.R.; Galea, V. Enhancing the ability of gait analyses to differentiate between groups: Scaling gait data to body size. *Gait Posture* **2001**, *13*, 193–201. [[CrossRef](#)]
33. Dubost, V.; Kressig, R.W.; Gonthier, R.; Herrmann, F.R.; Aminian, K.; Najafi, B.; Beauchet, O. Relationships between dual-task related changes in stride velocity and stride time variability in healthy older adults. *Hum. Mov. Sci.* **2006**, *25*, 372–382. [[CrossRef](#)]
34. Giakas, G.; Balzopoulos, V. Time and frequency domain analysis of ground reaction forces during walking: An investigation of variability and symmetry. *Gait Posture* **1992**, *5*, 189–197. [[CrossRef](#)]
35. Paul, D.; Su, R.; Romain, M.; Sébastien, V.; Pierre, V.; Isabelle, G. Feature selection for outcome prediction in oesophageal cancer using genetic algorithm and random forest classifier. *Comput. Med. Imaging Graph.* **2017**, *60*, 42–49. [[CrossRef](#)]
36. Li, H.; Yuan, D.; Ma, X.; Cui, D.; Cao, L. Genetic algorithm for the optimization of features and neural networks in ECG signals classification. *Sci. Rep.* **2017**, *7*, 41011. [[CrossRef](#)]
37. Gokulnath, C.B.; Shantharajah, S. An optimized feature selection based on genetic approach and support vector machine for heart disease. *Clust. Comput.* **2019**, *22*, 14777–14787. [[CrossRef](#)]
38. Snoek, B.J.; Larochelle, H.; Adams, R.P. Practical bayesian optimization of machine learning algorithms. *arXiv* **2012**, arXiv:1206.2944v2.
39. Kim, J.; Lee, G.; Heimgartner, R.; Revi, D.A.; Karavas, N.; Nathanson, D.; Galiana, I.; Eckert-Erdheim, A.; Murphy, P.; Perry, D.; et al. Reducing the metabolic rate of walking and running with a versatile, portable exosuit. *Science* **2019**, *365*, 668–672. [[CrossRef](#)]
40. Grefenstette, J. Optimization of Control Parameters for Genetic Algorithms. *IEEE Trans. Syst. Man Cybern.* **1986**, *16*, 122–128. [[CrossRef](#)]
41. Hasançebi, O.; Erbatur, F. Evaluation of crossover techniques in genetic algorithm based optimum structural design. *Comput. Struct.* **2000**, *78*, 435–448. [[CrossRef](#)]
42. Chicano, F.; Whitley, D.; Alba, E. Exact computation of the expectation surfaces for uniform crossover along with bit-flip mutation. *Theor. Comput. Sci.* **2014**, *545*, 76–93. [[CrossRef](#)]