Engineering Research
Express

PAPER

Automated classification of post-operative gait abnormalities following hip surgery using machine learning

To cite this article: Md Mohiuddin Soliman *et al* 2025 *Eng. Res. Express* **7** 035203

IOP Publishing Eng. Res. Express 7 (2025) 035203 https://doi.org/10.1088/2631-8695/ade84d

Engineering Research Express

CrossMark

RECEIVED
8 October 2024

REVISED
21 June 2025

ACCEPTED FOR PUBLICATION
25 June 2025

PUBLISHED
2 July 2025

PAPER

Automated classification of post-operative gait abnormalities following hip surgery using machine learning

Md Mohiuddin Soliman[1], Moajjem Hossain Chowdhury[1], M Murugappan[2,3,*] and Muhammad E H Chowdhury[4,*]

[1] Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia (UKM), Bangi 43600, Malaysia
[2] Intelligent Signal Processing (ISP) Research Lab, Department of Electronics and Communication Engineering, Kuwait College of Science and Technology, Block 4, Doha, 13133, Kuwait
[3] Department of Electronics and Communication Engineering, Vels Institute of Sciences, Technology, and Advanced Studies, Chennai, Tamil Nadu, India
[4] Department of Electrical Engineering, Qatar University, Doha 2713, Qatar
[*] Authors to whom any correspondence should be addressed.

E-mail: mchowdhury@qu.edu.qa **and** m.murugappan@kcst.edu.kw

Keywords: machine learning, gait cycle, human activity, hip surgery and hip replacement

Supplementary material for this article is available online

Abstract

An injury, chronic illness, obesity, infection, and more can negatively affect the hip joint. Surgery and implant placement are the standard treatments for moderate to severe hip issues. These treatments, however, can alter a patient's gait patterns. Gait patterns must be assessed clinically by a qualified physician and a specialized examination is required to detect and monitor these changes. By contrast, Machine Learning (ML) techniques assist in diagnosing a wide variety of anomalies and illnesses. In addition to being extremely accurate, it reduces subjectivity in clinical expert evaluations. Gait anomalies can also be quickly identified and monitored inexpensively and quickly using ML. Three open-source datasets (GaitRec, Gutenberg, and Orthoload) were utilized in this study for the gait cycle conditions for healthy control, hip surgery, and hip implant patients. This study classifies individuals into two classes: Healthy control/Gait Abnormality and three classes: Healthy control/ Hip surgery/ Hip implant by gait cycle conditions using only vertical ground reaction forces (vGRFs) from these datasets, which consist of 3D GRFs. The essential steps in data preparation include filtering, denoising, normalizing, resampling, and augmenting. The purpose of these efforts was to improve the model performance in classification and reduce biases. We used several feature extraction techniques, focusing on excluding highly correlated features. The final analysis utilized five widely recognized feature selection algorithms (Minimum Redundancy Maximum Relevance (mRMR), Neighborhood Component Analysis (NCA), Multi-Cluster Feature Selection (MCFS), Chi-square, and Relief) to arrange the features systematically. Based on a comprehensive examination of five machine learning classifiers (k-nearest neighbor (KNN), artificial neural network (ANN), decision tree (DT), support vector machine (SVM), and Naïve Bayes (NB)), The KNN classifier exhibited the highest level of accuracy. The two and three-class classification's overall accuracy, precision, sensitivity, and F1 score are 95.48%, 96.13%, 95.48%, 95.63% and 89.18%, 89.30%, 89.18, and 87.15%, respectively. With the proposed solution, clinicians can more easily identify gait abnormalities based on vertical ground reaction forces.

1. Introduction and related works

The pattern of limb movements a person makes while moving around is known as their gait, and it can be either natural or learned through particular training. Collectively, these styles are referred to as human gaits. Gait pattern analysis is a clinical tool for quantifying human movement and assessing gait performance. Gait pattern analysis can detect abnormalities in gait cycles caused by various factors, such as injuries, arthritis, congenital disabilities, neurological conditions (like cerebral palsy or stroke), hip joint issues, psychological disorders, etc [1]. As of 2022, 1,334,357 primary hip replacements were reported to the National Joint Registry (NJR), of which 30.7% were cemented, 37.1% were cement-less, 23.7% were hybrids, 2.6% were reverse, 3.1% were resurfacing, and 2.9% were unclassified [2]. The gait patterns of patients change following surgical procedures and implants compared to healthy individuals, necessitating their extensive observation and identification.

As a result of notable variations, complexity, non-linear interactions, and temporal complexities, clinical gait analysis generates extensive and complex data [3]. Due to the complexity and time-consuming nature of interpreting gait patterns, expert medical interpretation is necessary. The role of machine learning has been demonstrated to enhance diagnostics, therapy, medication research, surgeries, and patient monitoring, revolutionizing various industries [4–7]. Various automated gait analysis algorithms based on Machine Learning (ML) have emerged in recent years, assisting physicians in the recognition and classification of clinically notable gait patterns [3, 8]. There are several machine learning approaches commonly used in this field, including artificial neural networks (ANN), K-nearest neighbor classifiers (KNNs), Support Vector Machines (SVMs), and various clustering techniques [3, 8–13]. Despite this, the efficacy of these methods is heavily dependent on the representation of the input data. Gait assessment often depends on kinematic gait parameters, including temporal-distance parameters and local maxima and minima of gait signals [3, 8, 9, 11].

Researchers have used different approaches to classify patients with gait disorders. Even though some have addressed it as a two-class problem (distinguishing healthy individuals from those with gait disorders), others have sought to categorize gait disorder patients based on severity levels, creating three, four, and five-class classification problems. There have also been studies that combine both approaches. According to [3, 8–17], the researchers used ML and deep learning techniques to classify the gait cycles of healthy and gait-disordered subjects. Specifically, they use the Support Vector Machine (SVM), the K-Nearest Neighbors (KNN), the Random Forest (RF), the Decision Tree (DT), and the Artificial Neural Network (ANN) as classical approaches, together with Deep Learning techniques like Convolutional Neural Networks (CNN) and Recursive Neural Networks (RNN). In this case, the minimum and maximum accuracies are 83.3% and 100%, respectively, because binary classification makes it much easier to solve. Furthermore, authors in [18–24] categorize a particular gait disorder based on its severity level (low, mild, severe) or stage (such as Parkinson's Disease (PD), autism spectrum disorders (ASD), Huntington's Disease (HD), and post-stroke gait disorders). Using the same methods as previously mentioned. Those cases had minimum and maximum accuracy of 62% and 99.7%, respectively. The studies in [25–28] attempted both types of classification, recognizing healthy and gait disorders and stratifying the severity of the disorders. Zhang *et al* [27] developed an ML framework for gait classification of elderly, stroke survivors, and Huntington's disease patients based on inertial sensors. However, researchers in [29–31] utilized CNNs, classical machine learning algorithms to classify pathological gait patterns. Using a Microsoft Kinect sensor, a solitary case study was presented about the feasibility of tracking movement patterns before and after total hip replacement surgery, using an inexpensive and portable vision-based system [31]. Using this approach, the study aimed to determine whether it would be possible. In most existing research studies, IMU, 3D-IMU, and EMG signals are utilized for the classification of gait abnormalities using ML and DL methods. The existing systems lack generalizability due to a limited number of samples as well as being collected under constrained environmental conditions. It is more common for early works to consider binary classes as opposed to multi-class classifications. In this study, the results indicated that an automated algorithm and the Kinect can be integrated into a patient's residence. This will enable them to detect changes in mobility during recovery from an injury or illness.

Based on a thorough review of previous literature, the following limitations can be identified and summarized as follows: In the past, most of the investigations have typically relied on small datasets or with a limited number of samples for developing gait abnormality detection using GRF signals. However, detailed investigations on larger datasets are essential to ensuring the generalizability of classification results. In addition, the feasibility of incorporating the results of earlier studies directly into clinical practices is highly limited. In the case of training and testing the machine learning models, most of the previous studies used random stratification for training and testing, which is not always well-suited or robust for practical applications.

The limitations listed above have inspired and motivated the authors to conduct a more comprehensive investigation to fill in these gaps. In our earlier work [36], we utilized two different datasets [GaitRec and Gutenberg] to classify the gait disorders using different types of classical machine learning algorithms. Here, we have performed both binary and multi-class classification and achieved a mean accuracy over 95%. To the best of our knowledge, no previous study has attempted to distinguish between gait cycle abnormalities in patients requiring hip surgery and those receiving hip implants. As a result, this classification can be extremely useful for monitoring patient progress and making informed decisions based on their development. Using machine learning to facilitate these investigations is the primary contribution of this paper. The authors attempted to

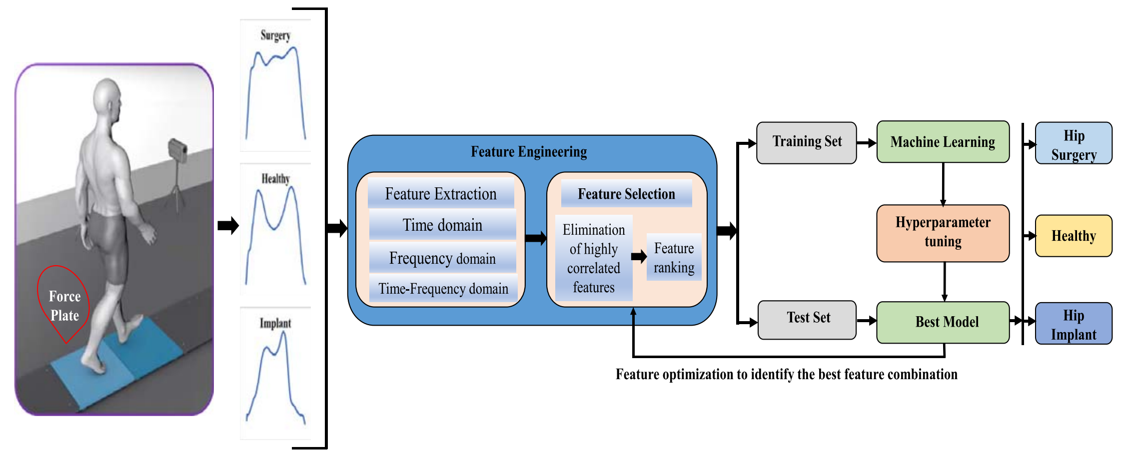

Figure 1. Overview of the methodology.

classify the gait patterns of three groups: healthy controls, patients who have had hip surgery without hip replacement, and patients with hip replacement. The classification was achieved using machine learning techniques and vertical ground reaction force (vGRF) signals. This study aims to present a system for assessing the health of patients, hip surgery patients, and hip-implanted patients from gait analysis rather than x-rays. Moreover, this will be useful for assessing the hip implant on a daily, weekly, or monthly basis in a home setting, allowing the patient to monitor their hip implant condition from home and visit the hospital in case of an abnormality. Thus, clinicians will be able to manage patients more effectively, and their workload will be reduced.

2. Materials and methods

Figure 1 describes the methodology employed in this study. We used three prevalent benchmark datasets to automate the prediction of abnormal gait patterns. Further details about these datasets are provided in the following section (see section 2.1.1). Vertical Ground Reaction Force (GRF) signals were acquired using embedded force plates from both right and left feet. This study used vGRF signals to extract features across time, frequency, and time-frequency domains to distinguish between healthy control (HC) subjects, those who underwent hip surgery without hip arthroplasty, and those with hip implants in binary and three-class classifications. Machine learning models were employed to perform various classification tasks on a large benchmark dataset, providing comprehensive insights into predicting joint-specific gait abnormalities and their implications.

vGRF signals are applied to the system for feature engineering (feature extraction and selection), and then the Dataset is divided into two subsets: training and test. The training dataset then trains several machine-learning models. The hyperparameters were tuned to improve performance, and the best model was found with the best combination of features, and the results were reported.

2.1. Patient's characteristics and data overview

Based on their previous medical diagnosis, a physical therapist puts each patient's information into a class by hand. Figure 2 shows the sample of vGRF signals of three different classes (control, hip implant, and hip surgery) considered in this work. In each subject, we have provided a detailed description:

- **Healthy control:** A public appointment system was used to recruit healthy individuals from the clinic's surrounding area. Participants in the study had to be free of any discomfort or objections relating to their lower extremities or spine. Furthermore, they were not permitted to wear orthotics or orthopaedic insoles.
- **Hip Surgery:** In this category, individuals with hip injuries such as thigh and pelvis fractures, hip dislocations, and coxarthrosis are included.
- **Hip Implantation:** Patients in this group had hip implant surgery involving total hip replacement via the direct lateral approach to treat primary hip osteoarthritis.

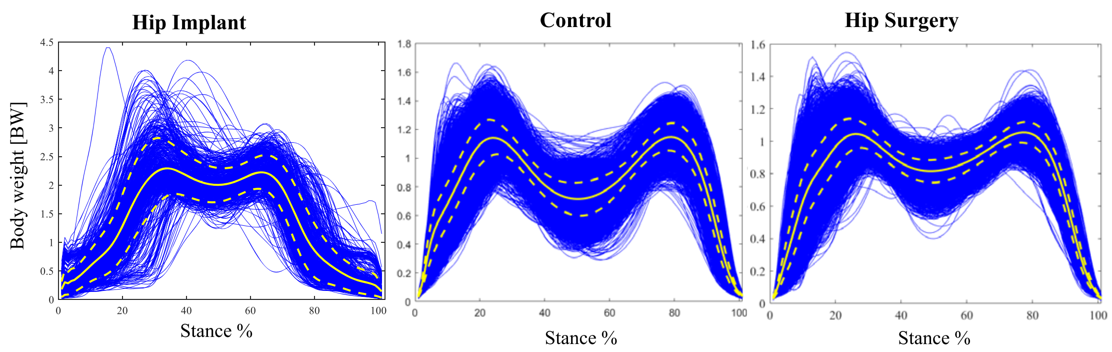

Figure 2. Visualizing means and standard deviation for each class (vGRF) as participants walked at self-selected speeds.

Table 1. Details demographic information of the total database. Reproduced from [30, 33, 35]. CC BY 4.0.

Database name	Category/class	Subject	Sample	Age (years)mean	Sex(m/f)	Body mass kg (Mean)
GaitRec	Hip surgery	450	12748	34.7	373/77	82.4
	Healthy control	211	2638	42.6	104/107	73.9
Gutenberg	Healthy control	350	8820	24.2	205/142	70.7
	Hip implant	17	594	61.5	14/5	76.9
Total		**1030**	**24800**	**40.7**	**696/331**	**75.9**

2.1.1. GaitRec and gutenberg database

In the gait classification challenge, we used a public dataset called GaitRec [32] and Gutenberg [30], published and maintained by an Austrian rehabilitation clinic. In this study, 450 patients out of 2085 underwent hip surgery, and 211 subjects were assigned to healthy controls. In contrast, the Gutenberg Gait Database [30] contains the gait data of 350 healthy controls collected over seven years. The subjects walked at a self-selected pace for two sequential walks while two force plates recorded ground reaction forces (GRFs) and center of pressure (COP). The data contributor on this Dataset has completed four unpublished investigations, and some of the data contributor's previously published data have been included in this Dataset [29, 30, 32]. A total of 350 healthy controls (205 males, 142 females, and 3 unknowns) between the ages of 11 and 64 participated in the study.

2.1.2. Orthoload database

In vivo measurements of orthopedic implant loads obtained via instrumented implants are stored in a publicly accessible repository. Load data are available for hip joints, shoulder joints, knee joints, vertebral body replacements, and internal spinal fixators. This study focuses on hip joint data collected from 19 patients with osteoarthritis and coxarthrosis using three different implants [33]. Data were collected for each patient under various loading conditions, but only level walking at a self-selected speed and walking on a treadmill at 4 km speed are merged here (compared with the other two datasets). Detailed information about this Dataset (Force F and Moment M) is available on their website [34]. Figure 2 shows an overview of the combined databases. This study considers data from these three public databases, as shown in table 1. As a result of combining both categories of the healthy control group from GaitRec and Gutenberg, three classes were calculated, which is the final data for this study.

2.2. Preprocessing

This Dataset included the center of pressure (COP) and three degrees of freedom analog ground reaction force (vGRF) signals, vertical, anterior-posterior (AP), and media-lateral (ML). Still, in this work, only vertical ground reaction force (vGRF) is considered as other datasets include only vGRF. We only considered vGRF in this study for several reasons. In gait analysis studies, vGRF alone provides enough information to assess gait. There is minimal contribution from other dimensions. Furthermore, 3D GRF is only possible using force plates. It is challenging to use force plates outside of a laboratory setting. Due to this, it is much more prudent to use vGRF, which can be obtained by various other GRF measurement methods, such as wearable insoles and inertial measurement units (IMU). Therefore, this work can be deployed with other technologies [36, 37].

A sampling frequency of 2 kHz was initially used to gather the signals. However, the raw signals were downscaled to 256 Hz for compatibility reasons aligned with the center's internal standards. The major reasons

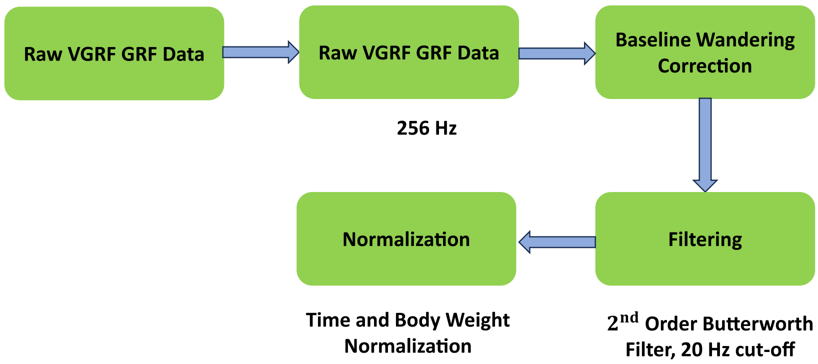

Figure 3. Overall data preparation steps.

for down sampling the vGRF signals to 256 Hz are: (i) most of the biomechanical signals have useful information to understand the gait abnormalities below 100 Hz. The lower sampling frequency of 256 Hz is sufficient to capture the minute changes in gait patterns and to reduce the effects of high frequency noises present in the signal for efficient gait pattern classification. (ii) Down sampling the vGRF signals to 256 Hz helps to reduce the computational complexity (time and memory) of gait pattern classification using Raspberry PI 3. (iii) The lower sampling frequency also helps to benchmark the performance of the model against different datasets. Thus, we worked with the available down-sampled version and further down-sampled the data to 256 Hz. In the next step, a 2nd order zero-lag Butterworth filter with a cut-off frequency of 20 Hz was applied. As shown in figure 3, the signals were first amplitude-normalized to 100% of body weight, then time-normalized to 101 samples, which correspond to the entirety of the stance phase. Different individuals have different body weights, which influences their vGRF signals. The vGRF force values would be affected by each subject's weight without normalization, making it difficult to distinguish biomechanical differences from simple weight differences. Biomechanical models and simulations assume forces relative to body weight. Indeed, researchers and clinicians can use normalization to reduce variance due to body weight differences. Hence, comparing these models directly is easier with normalization. By normalizing body weight, comparisons are fair regardless of mass. To avoid the influence of walking speed on the study result, the gait data was collected at a walking speed of 4 kph. Although many recommendations exist for an adequate cut-off frequency, 20 Hz appears to be a good compromise between minimizing noise and maximizing physiological frequency content.

2.3. Feature extraction

We extracted features from vGRF signals across several different domains, including time-domain (TD) [38], frequency domain (FD) [39, 40], and time-frequency domain (TFD) [41]. TD features are derived from the signal in the time domain by computing peak-to-peak values and other statistical features (such as kurtosis, skewness, interquartile range, etc) with pure TD characteristics (such as peak-to-peak values). TD features also include features obtained from the envelope of the signal such as the ratio between the maximum and minimum of the envelope, the time at which the envelope peaked, and its amplitude, etc shown in detail in Supplementary table 1. A time-domain feature accurately reflect inherent features of a time-domain signal. Indeed, the abnormal peak-to-peak values can indicate altered load-bearing patterns, often observed in individuals with joint impairments or post-surgical adaptations. The statistical features describe the distribution characteristics of the vGRF signal. For instance, increased kurtosis may suggest abrupt force applications, while skewness can reveal asymmetries in gait, potentially indicating compensatory mechanisms due to pain or weakness. There are 50 TD features that were considered in this work for gait abnormality detection (Supplementary table 1). Features that described the shape and size of the signal in the frequency domain representation were incorporated into FD features, including simple features such as the frequency with maximum power (dominant frequency) and more complex features such as the spectral roll-off point and spectral deformation.

In specific, dominant frequency features identify the primary frequency component of the vGRF signal. Deviations from typical dominant frequencies can reflect changes in gait cadence or the presence of pathological conditions affecting rhythmic movement patterns. The spectral roll-off point represents the frequency below which a certain percentage of the total spectral energy is contained. Shifts in this point can indicate alterations in gait dynamics, such as changes in step regularity or stability. Finally, the spectral deformation assesses deviations from expected spectral patterns, which can be associated with irregular force applications or neuromuscular control issues during gait. There are 24 FD features that were considered in this work for gait abnormality detection and are listed in Supplementary table 2. Based on discrete wavelet transforms (DWTs) [41, 42], TFD

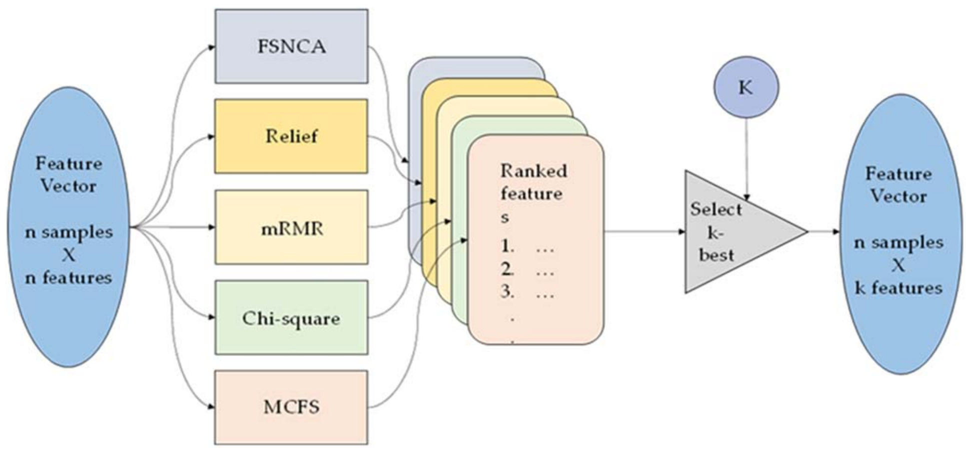

Figure 4. The feature selection process uses 5 algorithms: FSCNA, Relief, mRMR, Chi-square statistics, and MCFS ranking.

features include band power, mean absolute value, waveform length, root mean square value, standard deviation, and fractal length of extracted wavelets (see Supplementary table 3). These features, derived from discrete wavelet transforms, capture transient changes in the vGRF signal. They are sensitive to abrupt force variations and can detect subtle gait abnormalities, such as those arising from early-stage musculoskeletal disorders. By analyzing these features, our study aims to provide a comprehensive assessment of gait, facilitating the detection of post-operative gait abnormalities and contributing to improved patient outcomes. A total of 191 features were extracted, including 50 TD features (Supplementary table 1), 24 FD features (Supplementary table 2), and 117 TFD features (Supplementary table 3). MATLAB code was developed in-house, and these features were used to train and test classical machine learning algorithms.

Before doing the feature ranking, strongly correlated features were removed to avoid influencing the algorithm incorrectly. The correlation matrix between features is computed using pairwise linear correlation. Whenever a feature has a correlation coefficient (CC) of 0.9 or more, it is deemed heavily correlated, and one of those features is removed from the feature vector [53, 54]. Following the removal of highly correlated features in each experimental scheme, Supplementary figure S1 shows the number of features that remained.

2.4. Feature wselection

Even after removing highly correlated features, many features may still confound the model. Feature selection provides even more dimensionality reduction [43, 44]. A five-fold cross-validation technique was used in this present work to split the total number of features in five independent sets. Later, feature selection was applied to the training data on each fold, and the overall feature importance score was computed across all five folds. The features were then ranked based on their importance scores. This study employs a feature selection technique based on machine learning model, where one feature is added at a time, and its performance is evaluated. The feature matrix is reordered according to the significance of the features to ensure that the most eminent and useful features are utilized. We used MATLAB's selection of filter-type feature algorithms, including Minimum Redundancy Maximum Relevance (mRMR), Forward Selection Neighborhood Component Analysis (FSNCA), Multi-Cluster Feature Selection (MCFS), Chi-square, and Relief. The supplementary materials concisely explain their application. Figure 4 shows five different algorithms are used to select features: FSCNA, Relief, mRMR, Chi-square statistics, and MCFS ranking.

2.5. Machine learning

Out-of-fold data (4960) is used for testing (20% of the whole data), while the remaining data is divided into 80% and 20% as training (15872) and validation (3968) sets, respectively (table 2).

Based on GaitRec segments, Gutenberg, and sections of Orthoload databases, two classification approaches two-class classification (Healthy Control versus Gait Abnormality) and three-class classification (Healthy Control versus Hip Surgery versus Hip Implant). Two different experiments (classification without hyperparameter tuning and with hyperparameter tuning) were conducted in the three-class classification study. An experimental design employs subject-wise stratification for both two-class and three-class classifications. Following the processes of feature extraction and, where applicable, feature selection, various machine learning models were trained to regulate the highest-performing model among them, including Discriminant Analysis Classifier (DAC), Decision Trees (DT), Kernel Classification Models (KCM), K-Nearest Neighbor Models

Table 2. Description of the training, validation and testing samples used for each fold.

Training	Validation	Testing	Total
15872	3968	4960	24800

(KNN), Naive Bayes (NB), Support Vector Machines (SVM), and Artificial Neural Networks (ANNs). However, the five most promising classifiers (KNN, SVM, NB, DT, and ANN) were selected for this work based on their superior performance in classification and reported in the work [49].

A five-fold cross-validation technique was used to assess machine learning model performances. The optimal features from GaitRec, Guttenberg and orthoload datasets were divided into five folds, ensuring each subject's data appeared only once. The process was iterated five times, with four folds for training and one for testing. Over these five runs, the final performance results are presented as the average Accuracy, Sensitivity, Precision, and F1 score.

2.6. Hyperparameter tuning

Machine learning algorithms are tuned by optimizing their parameters to maximize their performance, a process known as hyperparameter tuning. Choosing optimal hyperparameters is essential to enhance model performance. There are several automated hyperparameter optimization techniques, but these are often problem-specific and differ from one problem to another. Typically, hyperparameter tuning involves changing the parameter's value and testing the model's performance to find the best value. To tune the hyperparameters of a machine learning algorithm, significant trial and error is required since each algorithm has several hyperparameters to adjust. In this study, this process is automated using Bayesian methods [48]. Taking a range of values for each hyperparameter, testtrains, and test each parameter iteratively. A selection of the optimal combination is made after examining all conceivable combinations. As part of the tuning procedure, the study examined the hyperparameters for each of the top four machine learning algorithms. Below are the parameters and values examined for these parameters.

- **Naive Bayes (NB):** NB implementation includes an additive smoothing parameter 'alpha' which varies between 0 and 1. The parameters examined were 0, 0.25, 0.5, 0.75, and 1.0.
- **SVM:** The regularization parameter 'C' and the kernel function were tuned for SVM. The values for 'C' were 1.0, 3.0, 5.0, and 10.0. The radial-basis function and the polynomial kernel were assessed for the kernel.
- **KNN:** In the KNN algorithm, the crucial adjustable parameter is the number of neighbors, denoted as 'k'. We experimented with several values for this parameter, including 1, 3, 5, 7, 15, 30, and 77, to determine the optimal choice for the algorithm.
- **DT:** The 'max depth' parameter of the DT classifier represents the maximum tree depth. This parameter was tested with values of 2, 3, 5, 10, 20, and 40.

2.7. Data augmentation

A model might introduce bias if trained on an unbalanced dataset due to the differences in signal quantities across classes. The use of data augmentation techniques on the training set is therefore crucial to achieving a uniform signal distribution across various classes, thus enhancing the credibility of model construction. The Synthetic Minority Over-Sampling Technique (SMOTE) was employed to balance the Dataset due to the varying sample sizes among classes [45]. Using this technique, minority class samples can be generated synthetically, mitigating imbalance issues and fostering an equal representation of classes, enhancing the model's accuracy.

2.8. Evaluation criteria

The performance of a classifier as a model is evaluated by five metrics. The terms true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP) are all used to measure diagnostic accuracy. These five metrics are outlined below:

Accuracy is defined as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

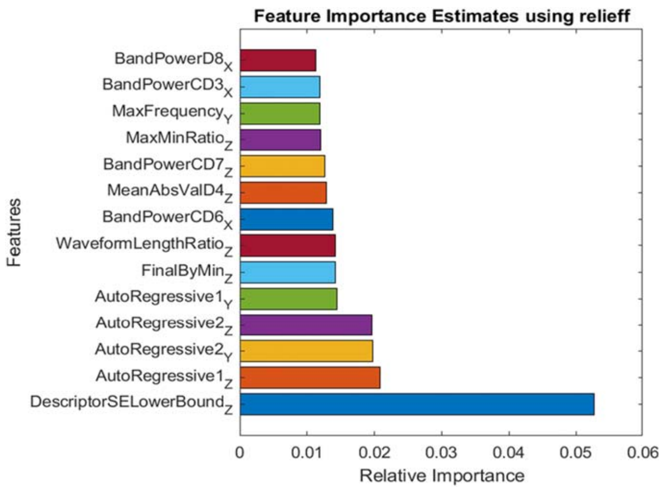

Figure 5. Top-ranked 14 features in classifying two-class classifications (Healthy Control versus Gait Abnormality).

Precision is defined by

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

Recall is defined by

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

F1-score is calculated as follows:

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{2 \quad \times TP}{2 \times TP + FN + FP} \tag{4}$$

Based on literature references, our primary measures were accuracy, sensitivity, specificity, and AUC [47]. Using the increasing feature search technique, we evaluated the model's classification performance across all experimental groups, progressively exploring combinations of top-ranked features to determine the optimal performance. To assess overall performance in multiclass scenarios, we analyzed individual Receiver Operating Characteristic (ROC) plots for each class and the macro average derived from the collective class ROC plots.

3. Numerical results and discussion

There are 1,030 subjects in this experiment, with 24,800 samples collected. A total of 1,030 participants participated in all the experiments, with equal numbers from each group being placed in the test and training sets. A variety of machine learning algorithms were tested, and the top five, which report higher performance, are presented here.

3.1. Two class classification

We reported two-class classification for Healthy Control versus Gait Abnormality with only the best performing feature selection algorithm (figure 5). Observing the feature importance plots (see figure 5) for two-class classification task, it can be seen that descriptor of signal entropy lower bound is the dominant feature in determining the Gait pattern of a subject. Along with it the autoregressive features also play an important role. Analyzing the important features in the aspect of channels, it can be observed that Z-axis GRF features contributed the most to the classification of the Gait pattern. This highlights how vital vertical GRF is for gait classification. However, as seen in figure 5, vertical GRF with the other two GRFs makes the model even more accurate. Figure 5 shows the top 14 features from the best feature ranking technique for the binary classification schemes. In figure 6, the ROC curves are plotted. The top 1 to the top 60 features are plotted for the binary classification task. This allows us to see the impact of different features on the model. Table 3 shows that for the two-class classification problem, KNN model is showing slightly better performance with Relief-based feature selection algorithm with an accuracy of 95.48%.

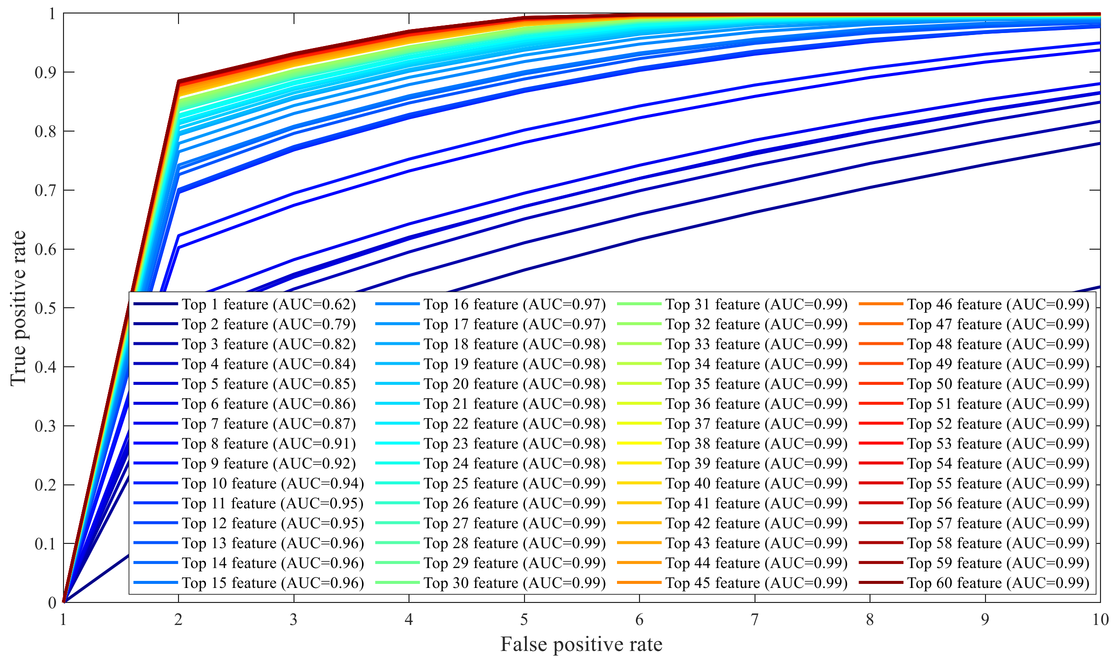

Figure 6. Two-class classification AUROC plot for healthy control versus gait abnormality.

Table 3. Healthy control versus gait abnormality classification performance.

Experimental schemes	Feature ranking technique	# of features	Accuracy(%)	Precision (%)	Sensitivity (%)	F1_Score (%)	AUC
	Relief	**Top 57**	**95.48**	**96.13**	**95.48**	**95.63**	**0.99**
HC/GD	mRMR	Top 58	93.24	94.56	93.24	93.54	0.99
	FSNCA	Top 48	93.73	94.89	93.73	93.99	0.99

3.2. Three class classification

We present the three-class classification for Healthy control versus Hip surgery versus Hip implant results based on subject-wise stratification. In subject stratification scenarios, the KNN classifier outperformed the other classifiers. Based on the results of a classification task, table 4 shows the results of five different classifiers. We have used KNNs, Naive Bayes, SVMs, DTs, and ANNs as classifiers. Classification features are the top 31, 30, 16, 44, and 42 features for KNN, Naive Bayes, SVM, DT, and ANN, respectively. We evaluated classifiers based on accuracy, precision, sensitivity, F1_Score, AUC, inference time, and model size. The SVM classifier had the highest average accuracy of 88.36%, followed by the KNN classifier at 87.08%, the DT classifier at 86.97%, the Naive Bayes classifier at 85.95%, and the ANN classifier at 84.83%. Among the five classifiers based on precision, the SVM scored the highest at 88.49%, followed by the KNN at 87.09%, the DT at 86.98%, the Naive Bayes classifier at 85.98%, and the ANN at 84.71%.

Based on sensitivity, the SVM classifier achieved 88.33%, followed by the KNN classifier at 87.08%, the DT classifier at 86.97%, the Naive Bayes classifier at 85.94%, and the ANN classifier at 84.87%. The SVM classifier achieved the highest F1_Score of 88.30%, followed by the KNN classifier at 87.08%, the DT classifier at 86.97%, the Naive Bayes classifier at 85.94%, and the ANN classifier at 84.87%. The SVM classifier was rated highest in terms of AUC with 0.950, followed by the KNN classifier at 0.950, the DT classifier at 0.950, the Naive Bayes classifier at 0.900, and the ANN classifier at 0.930. In terms of inference time, the Naive Bayes classifier had the fastest inference time of 4.36 μs, preceded by the SVM classifier with 0.91 μs, the DT classifier with 2.44 μs, the KNN classifier with 195.00 μs, and the ANN classifier with 340.80 μs. The ANN classifier had the largest model size of 9.49 MB, followed by the KNN classifier with 6.71 MB, the DT classifier with 6.64 MB, the Naive Bayes classifier with 6.55 MB, and the SVM classifier with 6.63 MB. Based on this task, the SVM classifier consistently outperformed the KNN classifier, the DT classifier, the Naive Bayes classifier, and the ANN classifier. An SVM classifier achieved the highest accuracy, precision, sensitivity, and F1_Score. With the Naive Bayes classifier, the inference time was the fastest, and the ANN classifier had the largest model size.

Table 4. Results of classification without hyperparameter tuning.

Classifier	Best FSA	Features num	Accuracy (%)	Precision(%)	Sensitivity(%)	F1_ score(%)	AUC	Stratification	Inference time (µs)	Model size (MB)
KNN	**Relief**	**Top 31features**	**87.08**	**87.09**	**87.08**	**87.08**	**0.95**	**Subject**	**195.00**	**6.71**
Naive Bayes	Relief	Top 30features	85.94	85.98	85.94	85.94	0.90	Subject	4.36	6.55
SVM	Relief	Top 16features	88.33	88.49	88.33	88.30	0.95	Subject	0.91	6.63
DT	Relief	Top 44 features	86.97	86.98	86.97	86.97	0.95	Subject	2.44	6.64
ANN	Relief	Top 42 features	84.87	84.71	84.87	84.87	0.93	Subject	340.80	9.49

Table 5. Results of classification with hyperparameter tuning.

Classifier	Best FSA	Features num	Accuracy (%)	Precision(%)	Sensitivity(%)	F1_ Score(%)	AUC	Stratification	Inference time (μs)	Model size (MB)
KNN	**MCFS**	**Top 34 features**	**89.18**	**89.30**	**89.18**	**87.15**	**0.96**	**Subject**	**198.20**	**6.74**
Naive Bayes	Relief	Top 30features	85.98	85.91	85.94	85.97	0.90	Subject	4.36	6.55
SVM	Relief	Top 16features	88.36	88.51	88.35	88.36	0.95	Subject	0.91	6.63
DT	Relief	Top 44 features	87.07	87.08	87.07	87.07	0.95	Subject	2.44	6.64
ANN	Relief	Top 42 features	84.88	84.73	84.88	84.89	0.93	Subject	340.80	9.49

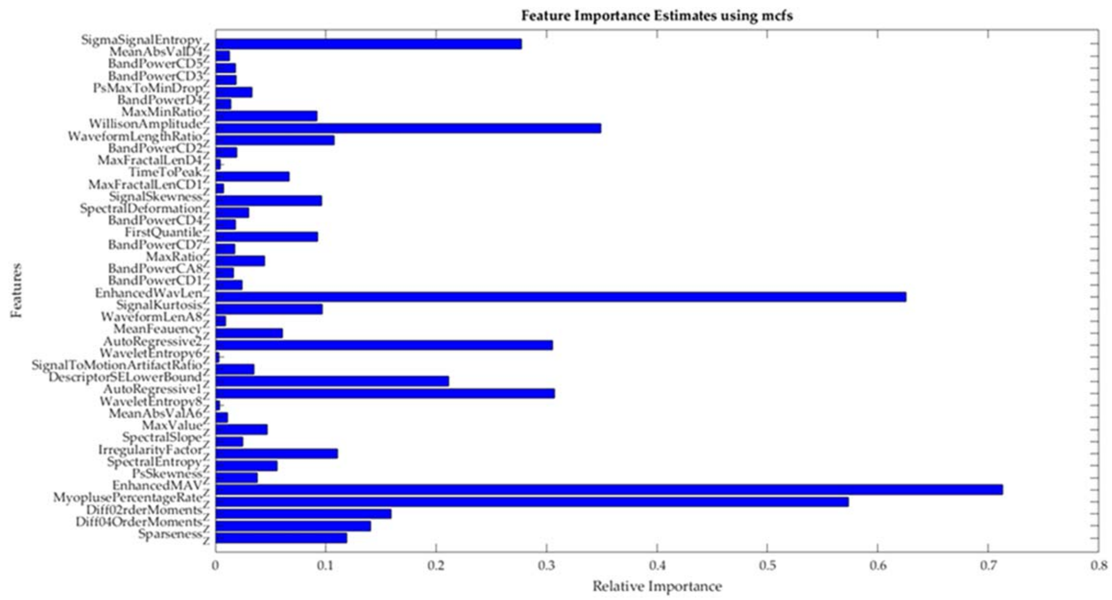

Figure 7. Top 34 features from the best feature ranking techniques for subject-wise stratification (Healthy control versus Hip surgery versus Hip implant) classification schemes.

The hyperparameter settings used in table 4 are default configurations and that no novel tuning methodology is applied. We also clarify that these results serve as a preliminary benchmark and are supplemented by further performance evaluation in subsequent sections where optimal hyperparameter configurations are applied as reported in section 2.6 and the improved results are reported in table 5.

We optimized the model hyperparameters, which improved the results slightly. For subject-wise stratification, the 'number of neighbors' parameter was set to 30 in KNN. In NB classifier, the optimal value of alpha that gives the highest mean classification rate is 0.25. In SVM, the radial-basis function (RBF) kernel with regularization parameter of 1.0 gives the highest mean classification rate in gait pattern classification. Finally, the maximum number of trees of 40 gives the highest accuracy in the decision tree classifier. In general, the classifiers achieved higher accuracy, precision, sensitivity, and F1_Score (table 5) using hyperparameter tuning. Table 4 shows the classifiers' performance without hyperparameter tuning. In addition, both tables had similar inference times and model sizes.

As for the feature selection algorithms used in these classifiers, for KNN we used MCFS, for Naive Bayes we used Relief, for SVM we used DT, and for ANN we used Relief. Accordingly, the top 34, 30, 16, 44, and 42 features are used for KNN, Naive Bayes, SVM, DT, and ANN, respectively. We evaluated the classifiers based on accuracy, precision, sensitivity, F1_Score, AUC, inference time, and model size. On average, the KNN classifier scored 89.18%, followed by the SVM classifier at 88.36%, the DT classifier at 87.07%, the Naive Bayes classifier at 85.97%, and the ANN classifier at 84.88%.

Figure 7 shows the top 34 features obtained through the best feature ranking techniques for subject-wise stratification of Healthy control versus Hip Surgery versus Hip Implant. According to the feature important plots, time domain features play a critical role in categorizing a subject's gait pattern. Each class's ROC plot was generated individually, and the macro average of these ROC plots was calculated. Each of these ROC plots is presented separately in figure 8. The importance of establishing a reliable clinical standard measure for diagnosing gait patterns cannot be overstated. A comprehensive review [46] of over 943 different classical techniques for analyzing gait patterns indicates that supervised Machine Learning algorithms have demonstrated accuracies exceeding 90% in gait analysis.

According to this study, the ML model developed in this study demonstrated high accuracy for random stratification cases and approached the standard range for subject-wise stratification. The achievement represents an encouraging first step toward automating the classification of gait abnormalities and establishes a baseline for future research. An extensive benchmark dataset comprising 1,030 subjects with 24,800 gait cycles was used for this study. Due to this, the developed model is expected to be able to generalize more efficiently. By extracting a substantial number (191) of features from time, frequency, and time-frequency domains from single channels, the model captures the variability within gait cycles, leading to better feature representations and enhanced learning. It is the first study to identify and distinguish gait cycle abnormalities among patients undergoing hip surgery or receiving hip implants. Further, it serves as a foundation for future research into

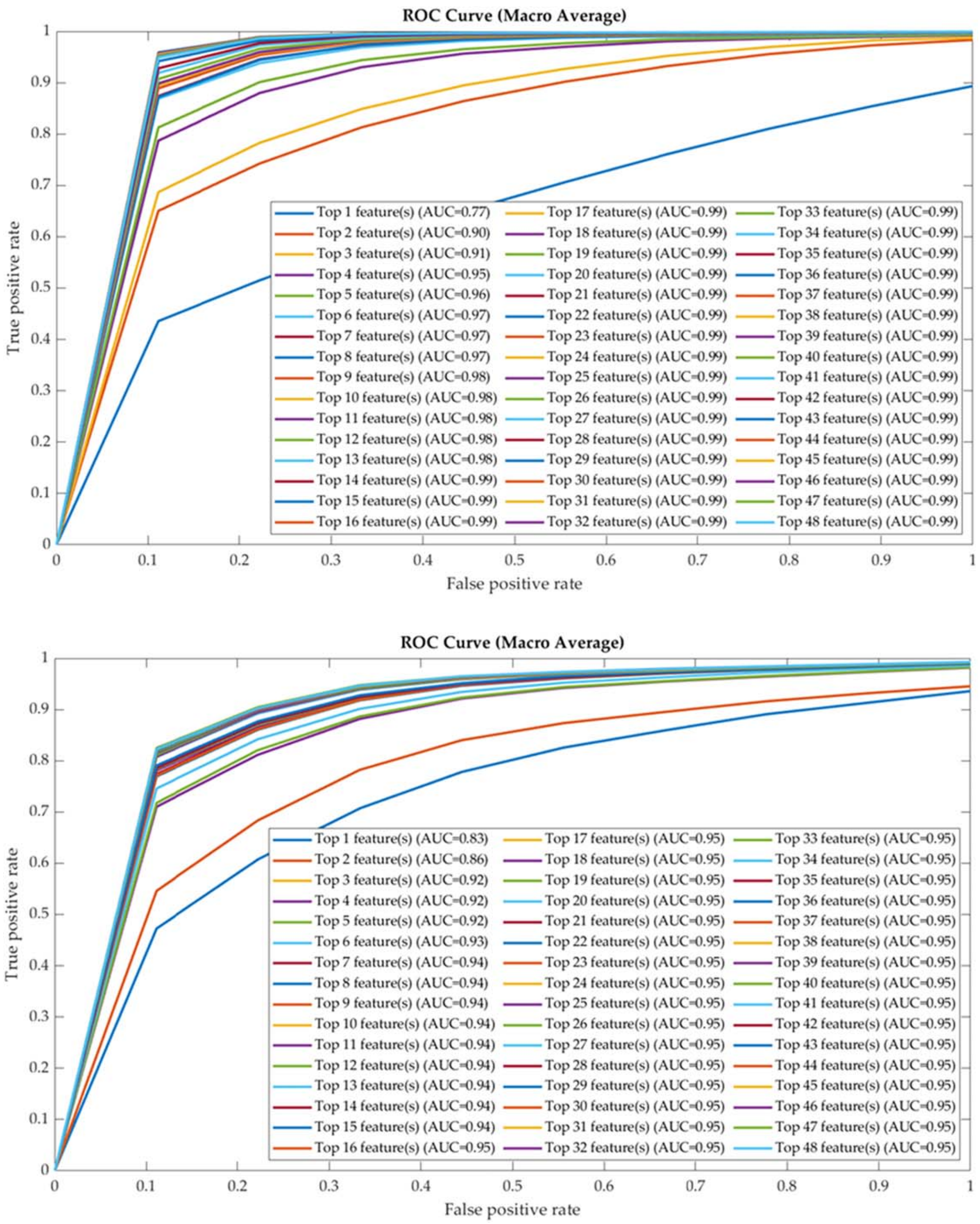

Figure 8. ROC plot for randon stratification (top) and subject-wise stratification (bottom).

categorizing Gait Cycle Disorder (GCD). As a result of experiments conducted across three datasets with various classification combinations, this study outperformed previous work in this area. This research focuses on classifying gait abnormalities post-hip surgery using vertical ground reaction forces (vGRFs), understanding joint interaction dynamics is crucial. Integrating models like the Hénon map could enhance our feature extraction process by accounting for joint synergies, potentially improving classification accuracy [50]. Furthermore, the methodology of analyzing muscle synergies to assess recovery aligns with our objective of monitoring post-operative gait [51]. Incorporating muscle synergy patterns as features in our machine learning models could provide a deeper understanding of neuromuscular coordination post-surgery [50, 51]. Monitoring muscle fatigue is vital in post-operative rehabilitation [52]. Applying similar synergy pattern analysis to lower limb muscles could enhance our ability to detect early signs of fatigue-related gait abnormalities.

In our future work, we aim to develop highly accurate deep neural networks for diagnosing diverse gait disorders, resulting in significant performance improvements using Convolutional Neural Network (CNN), Long Short Term Memory (LSTM), and Recurrent Neural Network (RNN). Additionally, adding a Generative

Table 6. Performance comparison of the present work with state-of-the-art models.

Authors	Goal	Participants	Best classifiers & performance
Zeng *et al* [18]	Gait pattern classification in Parkinson's disease patients and healthy controls	93 Parkinson's Disease patients and 73 healthy subjects	RBF Neural Network (2 class)Accuracy: 98.8%
Khandoker *et al* [25]	Automatic detection of gait patterns associated with balance deficiencies	13 healthy adults and 10 subjects with a history of falls	SVMAccuracy: 100%
Zhang *et al* [27]	Gait pattern recognition during lower extremity muscle fatigue and no-fatigue	17 healthy subjects	SVM (2 class)Accuracy: 96%
Slijepcevic *et al* [26]	Automated human gait analysis	182 healthy subjects and 728 different GD (4 types) patients	SVM (5 class)Accuracy: 62%
Alaqtash *et al* [9]	Automatic classification of pathological gait patterns from healthy walking	12 healthy volunteers, 4 CP patients, and 4 ME patients	KNNAccuracy: 95%
Wu *et al* [13]	Gait pattern categorization using kernel PCA feature extraction	24 young and 24 elderly	SVM (2 class)Accuracy: 91%
Wu *et al* [17]	classify gait patterns based on kinetic data	30 young and 30 elderly participants	SVM (2 class)Accuracy: 90%
Pogorelc *et al* [23]	A system to early detect abnormal gait patterns	5 healthy and 9 pathological elderly subjects	KNN (5 class)Accuracy: 99%
Kaczmarczyk *et al* [21]	Post-stroke patients' gait patterns classification	74 stroke survivors	ANN (3 class)Accuracy: 100%
Begg and Kamruzzaman [19]	Gait pattern classification in young and older individuals	12 young and 12 old participants	SVM (2 class)Accuracy: 100%
Chan *et al* [20]	Classification of younger and older people's gait patterns	13 healthy younger and 12 healthy elderly participants	MLP (2 class)Accuracy: 96.8%
Khalil *et al* [53]	Enhanced Binary Classification of Gait Disorders Using a Machine Learning Majority Voting Approach	2435 subjects	Accuracy of 96.63% in binary classification
Slijepcevic *et al* [54]	Decoding Gait Signatures: Exploring Individual Patterns in Pathological Gait using Explainable AI	2,092 individuals, including 1,283 cases with pathological gait	Classification accuracy of up to 99.3% (achieved with bilateral GRF3D data
Proposed Work	**Automated Classification of Post-Operative Gait Abnormalities following Hip Surgery using Machine Learning**	**561 healthy, 450 hip surgery, and 17 hip implants (total 1030) subjects**	**KNN (2-class)Accuracy: 95.48% for subject-wise stratification**
			KNN (3 class)Accuracy: 94.68% for Random stratification and 89.18% for subject-wise stratification

Adversarial Network (GAN) model to the training data could improve the model's performance. This method integrates synthetic and authentic data during training to equalize class distributions. Using such a framework can be immensely valuable for rehabilitation centers, allowing clinicians to meticulously monitor a patient's progress until complete recovery is achieved. Table 6 shows the comparison of performance of the proposed method with the state-of-the-art models used for gait abnormalities detection. Among the different works, this present work has utilized higher number of subjects (1030 subjects) and considered a maximum of three classes (healthy control, hip implant, and hip surgery). Based on the randon stratification method, the proposed method achieved a maximum mean classification rate of 94.68% and 89.18% on subject-wise stratification for three-class classification and 95.48% for a two-class classification.

To enhance the relevance and fairness of our comparative analysis, we have updated table 6 to include recent studies that utilized the same datasets. It is noteworthy that Khalil *et al* (2024) proposed a machine learning-based majority voting approach using the merged GaitRec and Gutenberg datasets, reporting strong binary classification results for gait disorder detection [53]. Additionally, Slijepcevic *et al* (2024) [54] explored explainable AI methods for class-specific gait analysis using GaitRec, demonstrating the utility of interpretable models in clinical gait evaluations. As a result of these comparisons, our multi-class classification approach reinforces its significance and validates its robustness against recent benchmarks based on similar data sets. The

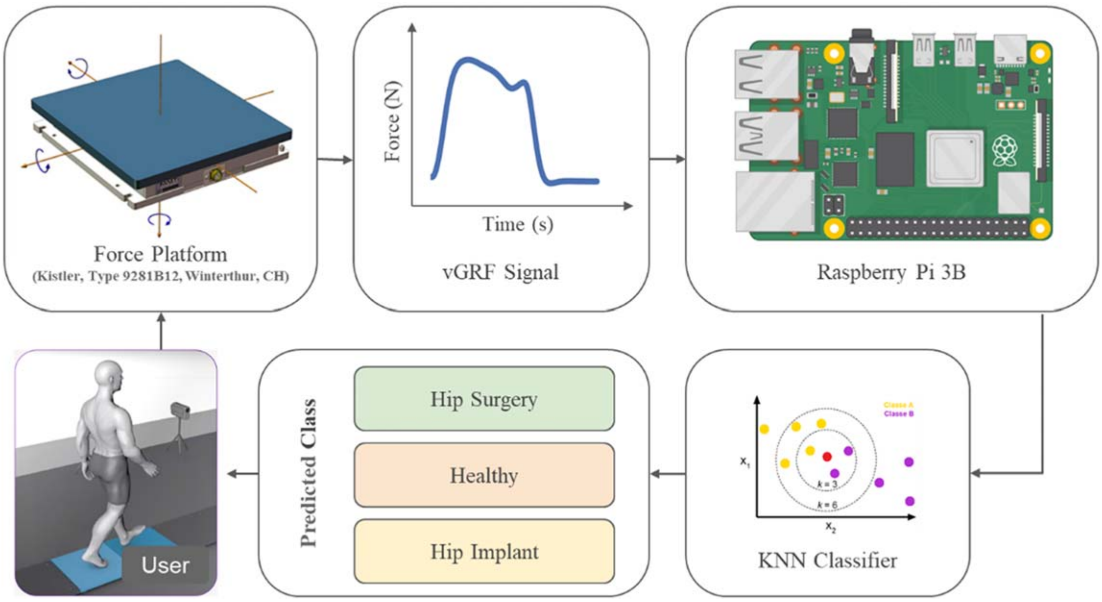

Figure 9. Real time implementation workflow.

results from these studies are not completely comparable with ours since none of them mentioned subject-wise studies.

4. Real-time implementation

A summary of the real-time implementation methodology can be found in figure 9. A vertical GRF signal was collected from a force plate for this study. Raspberry Pi 3Bs are used to control the circuitry, and ON/OFF switches are used to switch the device on and off manually. In terms of power consumption, the Raspberry Pi consumes about 3.45 Watts. Using Raspberry Pi, we have extracted features from the vGRF signal. The trained KNN classifier predicts predictably healthy controls (HC), Hip Surgery, and Hip Implanted using the extracted features. There is approximately 0.47-second inference time for each signal.

5. Conclusion

This study used three publicly available GRF resources: Gaitrec, Gutenberg, and specific segments from Orthoload. We homogenized the Gutenberg Gait Database with the GaitRec dataset to ensure their indistinguishability. This present work only utilizes 1D-GRF signals for gait abnormality classification. To accurately classify the gait patterns, different preprocessing steps such as filtering, denoising, normalizing, resampling, and augmenting data to reduce bias and make the training accurate. The features were extracted using different feature extraction techniques, the highly uncorrelated features were removed, and then the features were ranked according to five feature selection algorithms. Additionally, we tested several ML models, and the KNN model showed the best accuracy. Two-class subject-wise classification accuracy is 95.48% while subject-wise and random stratification of the three-class classification yield 89.18% and 94.68% accuracy, respectively. Finally, the KNN 3-class classification model is implemented in Raspberry PI 3B for real-time gait pattern classification, and it requires 0.47 sec inference time for each signal. It is the first study to recognize and differentiate between gait cycle abnormalities in patients undergoing hip surgery and those with implanted hips. We used classical machine learning algorithms to classify vertical ground reaction force (vGRF) signals.

Institutional review board statement

Not applicable.

Informed consent statement

Not Applicable.

Conflicts of interest

The authors declare no conflict of interest.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/Shuzan4/Hip_Implant-from-vGRF-.

ORCID iDs

Md Mohiuddin Soliman https://orcid.org/0000-0002-0268-9978
M Murugappan https://orcid.org/0000-0002-5839-4589
Muhammad E H Chowdhury https://orcid.org/0000-0003-0744-8206

References

[1] Stanhope S 2013 Measuring walking: a handbook of clinical gait analysis: Richard Baker, Mac Keith Press, 6 Market Road, London, N7 9PW, 2013, Paperback, 246*J. Biomech.;* **46** 2848
[2] Ben-Shlomo Y *et al* 2022 The national joint registry 19th *Annual Report* (https://ncbi.nlm.nih.gov/books/NBK587525/)
[3] Surmacz K, Redfern R E, Van Andel D C and Kamath A F 2024 Machine learning model identifies patient gait speed throughout the episode of care, generating notifications for clinician evaluation *Gait Posture.* (https://doi.org/10.1016/j.gaitpost.2024.09.001)
[4] Belyaev M, Murugappan M, Velichko A and Korzun D 2023 Entropy-based machine learning model for fast diagnosis and monitoring of parkinson's disease *Sensors* **23** 8609
[5] Yuvaraj R, Chadha S, Prince A A, Murugappan M, Islam M S B, Sumon M S I and Chowdhury M E H 2024 A Machine learning framework for classroom eeg recording classification: unveiling learning-style patterns *Algorithms.* **17** 503
[6] Johayra P *et al* 2025 A novel classical machine learning framework for early sepsis prediction using electronic health record data from ICU patients *Comput. Biol. Med.* **184** 109284
[7] Karthikeyan P, Murugappan M and Yaacob S 2012 EMG signal based human stress level classification using wavelet packet transform *Trends in Intelligent Robotics, Automation, and Manufacturing. IRAM* Vol 330 ed S G Ponnambalam, J Parkkinen and K C Ramanathan (Springer) Communications in Computer and Information Science (https://doi.org/10.1007/978-3-642-35197-6_26)
[8] Alharthi A S 2024 Interpretable machine learning comprehensive human gait deterioration analysis *Frontiers in Neuroinformatics* **18**
[9] Alaqtash M, Sarkodie-Gyan T, Yu H, Fuentes O, Brower R and Abdelgawad A 2011 Automatic classification of pathological gait patterns using ground reaction forces and machine learning algorithms *Annu Int. Conf. IEEE Eng. Med. Biol. Soc.* **2011** 453–7
[10] Hwang S, Kim J, Yang S, Moon H-J, Cho K-H, Youn I, Sung J-K and Han S 2024 Machine learning based abnormal gait classification with imu considering joint impairment *Sensors* **24** 5571
[11] Bogaert S, Davis J and Vanwanseele B 2024 Predicting vertical ground reaction force characteristics during running with machine learning *Front. Bioeng. Biotechnol.* **12** 1440033
[12] Mezghani N *et al* 2008 Automatic classification of asymptomatic and osteoarthritis knee gait patterns using kinematic data features and the nearest neighbor classifier *IEEE Trans. Biomed. Eng.* **55** 1230–2
[13] Wu J and Wang J 2008 PCA-based SVM for automatic recognition of gait patterns *J Appl. Biomech.* **24** 83–7
[14] Bengio Y, Courville A and Vincent P 2013 Representation learning: a review and new perspectives *IEEE Trans. Pattern Anal. Mach. Intell.* **35** 1798–828
[15] Giakas G and Baltzopoulos V 1997 Time and frequency domain analysis of ground reaction forces during walking: an investigation of variability and symmetry *Gait & Posture.* **5** 189–97
[16] Lafuente R, Belda J M, Sánchez-Lacuesta J, Soler C and Prat J 1998 Design and test of neural networks and statistical classifiers in computer-aided movement analysis: a case study on gait analysis *Clin. Biomech.* **13** 216–29
[17] Wu J, Wang J and Liu L 2007 Feature extraction via KPCA for classification of gait patterns *Hum. Mov. Sci.* **26** 393–411
[18] Zeng W, Liu F, Wang Q, Wang Y, Ma L and Zhang Y 2016 Parkinson's disease classification using gait analysis via deterministic learning *Neurosci. Lett.* **633** 268–78
[19] Begg R and Kamruzzaman J 2005 A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data *J. Biomech.* **38** 401–8
[20] Chan H *et al* 2013 Assessing gait patterns of healthy adults climbing stairs employing machine learning techniques *Int. J. Intell. Syst.* **28** 257–70
[21] Kaczmarczyk K, Wit A, Krawczyk M and Zaborski J 2009 Gait classification in post-stroke patients using artificial neural networks *Gait Posture.* **30** 207–10
[22] LeMoyne R, Kerr W, Mastroianni T and Hessel A 2014 Implementation of machine learning for classifying hemiplegic gait disparity through use of a force plate *Proc. 2014 13th Int. Conf. on Machine Learning and Applications* 379–82
[23] Pogorelc B, Bosnić Z and Gams M 2012 Automatic recognition of gait-related health problems in the elderly using machine learning *Multimedia Tools Appl.* **58** 333–54
[24] Soares D P, de Castro M P, Mendes E A and Machado L 2016 Principal component analysis in ground reaction forces and center of pressure gait waveforms of people with transfemoral amputation *Prosthet. Orthot. Int.* **40** 729–38

[25] Khandoker A H, Lai D T H, Begg R K and Palaniswami M 2007 Wavelet-based feature extraction for support vector machines for screening balance impairments in the elderly *IEEE Trans. Neural Syst. Rehabil. Eng.* **15** 587–97
[26] Slijepcevic D *et al* 2017 Automatic classification of functional gait disorders *IEEE Journal of Biomedical and Health Informatics* **22** 1653–61
[27] Zhang J, Lockhart T E and Soangra R 2014 Classifying lower extremity muscle fatigue during walking using machine learning and inertial sensors *Ann. Biomed. Eng.* **42** 600–12
[28] Burdack J, Horst F, Giesselbach S, Hassan I, Daffner S and Schöllhorn W I 2020 Systematic comparison of the influence of different data preprocessing methods on the performance of gait classifications using machine learning *Front. Bioeng. Biotechnol.* **8**
[29] Horst F, Eekhoff A, Newell K M and Schöllhorn W I 2017 Intra-individual gait patterns across different time-scales as revealed by means of a supervised learning model using kernel-based discriminant regression *PLoS One* **12** e0179738
[30] Horst F, Slijepcevic D, Simak M and Schöllhorn W I 2021 Gutenberg gait database, a ground reaction force database of level overground walking in healthy individuals *Scientific Data.* **8** 232
[31] Dolatabadi E, Taati B and Mihailidis A 2014 Vision-based approach for long-term mobility monitoring: single case study following total hip replacement *J. Rehabil Res. Dev.* **51** 1165–76
[32] Horst F *et al* 2016 Daily changes of individual gait patterns identified by means of support vector machines *Gait Posture.* **49** 309–14
[33] Damm P, Bender A, Waldheim V, Winkler T and Duda G N 2021 Surgical cup placement affects the heating up of total joint hip replacements *Sci. Rep.* **11** 15851
[34] Bergmann G, Bender A, Dymke J, Duda G and Damm P 2023 *Standard Loads Hip Joint* https://orthoload.com/test-loads/standardized-loads-acting-at-hip-implants/
[35] Horsak B, Slijepcevic D, Raberger A-M, Schwab C, Worisch M and Zeppelzauer M 2020 GaitRec, a large-scale ground reaction force dataset of healthy and impaired gait *Scientific Data* **7** 143 https://nature.com/articles/s41597-020-0481-z
[36] Nazmul Islam Shuzan M *et al* 2023 Machine learning-based classification of healthy and impaired gaits using 3D-GRF signals *Biomed. Signal Process. Control* **81** 104448
[37] Mai P and Willwacher S 2019 Effects of low-pass filter combinations on lower extremity joint moments in distance running *J. Biomech.* **95** 109311
[38] Altin C and Er O 2016 Comparison of different time and frequency domain feature extraction methods on elbow gesture's EMG *European Journal of Interdisciplinary Studies.* **5** 35
[39] El Halabi R 2018 Blind source separation of single-sensor recordings: application to ground reaction force signals *Physics* Université de Lyon English
[40] Stergiou N, Giakas G, Byrne J B and Pomeroy V 2002 Frequency domain characteristics of ground reaction forces during walking of young and elderly females *Clin Biomech (Bristol, Avon).* **17** 615–7
[41] Doulah A B, Fattah S A, Zhu W P and Ahmad M O 2014 DCT domain feature extraction scheme based on motor unit action potential of EMG signal for neuromuscular disease classification *Healthc Technol Lett.* **1** 26–31
[42] Soubra R, Diab M O and Moslem B 2017 A mother wavelet selection study for vertical ground reaction force signals *In: Proc. 2017 2nd Int. Conf. on Bio-engineering for Smart Technologies (BioSMART)* 1–5
[43] Haque F *et al* 2021 A nomogram-based diabetic sensorimotor polyneuropathy severity prediction using Michigan neuropathy screening instrumentations *Comput. Biol. Med.* **139** 104954
[44] Shuzan M N I *et al* 2021 A novel non-invasive estimation of respiration rate from motion corrupted photoplethysmograph signal using machine learning model *IEEE Access.* **9** 96775–90
[45] Chawla N, Bowyer K, Hall L and Kegelmeyer W 2002 SMOTE: synthetic minority over-sampling technique *J. Artif. Intell. Res. (JAIR* **16** 321–57
[46] Nascimento D H A, Magalhães F A, Sabino G S, Resende R A, Duarte M L M and Vimieiro C B S 2022 Development of a human motion analysis system based on sensorized insoles and machine learning algorithms for gait evaluation *Inventions.* **7** 98
[47] Hakkak Moghadam Torbati A, Pellegrino S, Fonti R, Morra R, De Placido S and Del Vecchio S 2024 Machine learning and texture analysis of [18f]fdg pet/ct images for the prediction of distant metastases in non-small-cell lung cancer patients *Biomedicines.* **12** 472
[48] Choi W, Yang W, Na J, Lee G and Nam W 2021 Feature optimization for gait phase estimation with a genetic algorithm and bayesian optimization *Applied Sciences* **11** 8940
[49] Haocong R, Minlin Z, Xuejiao Z and Chunyan M 2025 A survey of artificial intelligence in gait-based neurodegenerative disease diagnosis *Neurocomputing* **626**
[50] Torbati A H, Jami S and Kobravi H R 2022 Is the Hénon map able to predict the interaction dynamics between the knee and hip joints emerged during sit-to-stand movement? *Biomed. Phys. Eng. Express* **8** 045003
[51] Torbati A H, Abbasnezhad L and Tahami E 2017 Determination of the best recovery based on muscles synergy patterns and lactic acid *Journal of Human Sport and Exercise* **12** 180–91
[52] Torbati A H, Tahami E and Kobravi H R 2018 May 14 Clinical model to the analysis of synergy pattern changes of back muscles and its relationship with the occurrence of fatigue *The Open Bioinformatics Journal.* **11**
[53] Khalil A, Saad M, Chaar K, Tafreshi R, Abdulla S and Wahid M F 2024, Enhanced binary classification of gait disorders using a machine learning majority voting approach *2024 46th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE) 1–4
[54] Slijepcevic D, Horst F, Simak M L, Schöllhorn W I, Horsak B and Zeppelzauer M 2024 Decoding gait signatures: exploring individual patterns in pathological gait using explainable ai *IEEE Access* (https://doi.org/10.1109/ACCESS.2024.3513893)