WalkEar: Holistic Gait Monitoring using Earables

Jake Stuchbury-Wass*, Yang Liu*, Kayla-Jade Butkow*, Josh Carter†, Qiang Yang*, Mathias Ciliberto*,
Ezio Preatoni†, Dong Ma‡ and Cecilia Mascolo*
*University of Cambridge, UK, †University of Bath, UK, ‡Singapore Management University, Singapore
{js2372, yl868, kjb85}@cam.ac.uk, jc2369@bath.ac.uk, {qy258, mc2514}@cam.ac.uk,
ep314@bath.ac.uk, dongma@smu.edu.sg, cm542@cam.ac.uk

***Abstract*—Gait behaviour is a key health metric. Temporal, spatial and kinetic walking gait parameters are valuable in enhancing sport performance and early health diagnostics Full gait assessment requires a gait clinic and existing wearable gait tracking systems typically measure isolated subsets of parameters tailored to specific applications. This is useful when the condition to be monitored is known, but fails to offer a comprehensive view of an individual's gait traits when their pathology is unknown or changing, or a general assessment is required. To support holistic walking gait tracking, we introduce WalkEar, a novel sensing platform designed to simultaneously track gait parameters using commodity earbuds. WalkEar operates by detecting gait events to derive temporal gait parameters and segment the IMU data. WalkEar then progresses earable gait assessment by, for the first time, estimating kinetic gait parameters and reconstructing the vGRF curve using machine learning. Each parameter is calculated on a step-to-step basis for gait variability and asymmetry. We developed an earbud prototype and collected data from 13 participants using gold standard force plates and instrumented treadmill ground truth. Extensive experiments demonstrate the promising performance of WalkEar, achieving an overall MAPE of 5.1% in estimating gait, 2.0% MAPE on kinetic gait parameters, and an NRMSE of 5.3% for vGRF curve reconstruction.**

***Index Terms*—Wearables, Earables, Gait, Spatiotemporal gait parameters, Kinetic gait parameters, Ground reaction force.**

I. Introduction

Gait analysis, the study of human locomotion, is crucial for clinical evaluations and wellness monitoring. Gait analysis has wide-ranging applications in rehabilitation, sports science, fitness monitoring [1], and clinical settings, aiding in injury detection, gait disorder identification, and early diagnosis of conditions like Parkinson's disease [2]. Walking gait parameters can be categorized into *temporal* (e.g., cadence, stride, stance, and swing time), *spatial* (e.g., step length and vertical displacement of the center of mass), and *kinetic* (e.g., ground reaction forces, joint torques). While temporal and spatial parameters are commonly assessed and associated with overall fitness levels [1] and neurological conditions [3], kinetic parameters offer additional insight into biomechanical determinants of movement and the loading of anatomical structures [4] and pathologies such as osteoarthritis [5]. Additionally, the left to right asymmetries in gait, typically tracked in left-right differences in gait parameters, are commonly used to track and assess clinical disorders [6].

Outside of the lab, wearables are the main method for gait assessment. Current methods for wearable gait analysis involves attaching IMU sensing devices to the shoes [7], [8], ankles and lower legs [8], [9] or hips and lower back [9], [10]. However, these devices are often not widely available, are expensive devices with only a singular use case and can be socially unacceptable or unconformable to wear. To address these issues, some researchers have explored using commercial wearables like smartwatches [11]. Additionally, these methods typically focus on specific subsets of gait parameters for particular applications, lacking a versatile monitoring tool that benefits a broad population through ubiquitous wearables, where a pathology may also be unknown.Moreover, smartwatches suffer from strong interference due to the movement of human arms, yielding mediocre gait monitoring performance [12].

To fill this gap, we present WalkEar, an earable-based walking gait monitoring system able to simultaneously measure temporal, spatial, and kinetic parameters using commonly accepted wearable devices. WalkEar leverages IMUs in earphones to detect gait events and estimate parameters on a step-by-step basis. It employs lightweight techniques for efficient operation on mobile devices, compensates for different earphone orientations, and mitigates the impact of head movements. Compared to other wearable form factors, earables are desirable as they are widely available and have dual elements that are located centrally and symmetrically in the body, allowing symmetric monitoring of both feet. Additionally, the human head exhibits high stability during walking, which incurs limited interference and thereby enables more accurate gait monitoring.

While related works already estimate gait events for temporal parameters from earables [13], [14], [15], WalkEar requires these gait events for data segmentation to estimate the kinetic gait parameters. The WalkEar algorithm outperforms the related work for these parameters and enables the estimation of novel parameters. For kinetic gait parameters, WalkEar not only estimates useful scalar parameters such as weight acceptance force and loading rate but also reconstructs a continuous vGRF curve using a sequence-to-sequence model. This offers advantages over related work that are essential to comprehensive gait analysis, such as identifying subtle anomalies in force distribution and timing. WalkEar uses each of these parameters to estimate the user's left to right gait asymmetries.

We prototyped WalkEar using custom earphones equipped with 6-axis IMUs (3-axis accelerometer and 3-axis gyroscope). Data was collected from 13 participants who walked in various experimental settings, including lab-controlled conditions on an instrumented treadmill under different walking speeds, in-the-wild evaluation including free-walking on force plates, and

a stop-and-go-scenario validation resulting in over 18,000 step samples. WalkEar achieves an overall Mean Absolute Percentage Error (MAPE) of 5.1% for spatio-temporal parameters and 2.0% for kinetic parameters, as well as a Normalised Root Mean Square Error (NRMSE) of 5.3% on vGRF curve reconstruction. Compared to existing earable-based gait tracking systems, WalkEar not only expands the coverage of gait parameters but also delivers more accurate estimation with better generalization ability, validated by leave-one-subject-out evaluations.

The key contributions of this work include:

- Simultaneous measurement of temporal, spatial, and kinetic parameters, as well as their asymmetries, for holistic gait analysis in real-time, locally, on a smartphone.
- Designing lightweight signal processing methods to improve gait estimation results and enable the full WalkEar pipeline.
- Showing for the first time vGRF curve reconstruction from an earable device and estimating additional kinetic gait parameters.

II. Walking Gait Primer

A. Spatio-temporal gait parameters

Human gait involves two key events: heel strike (HS) and toe off (TO), heel strike refers to the moment when the heel contacts the ground and toe off occurs when the toes leave the ground, marking the point at which the foot is no longer in contact with the surface.

As illustrated in Figure 1, by connecting the two events of the two feet and subtracting the relevant timestamps, four temporal parameters can be defined: ***Cadence***, ***Stride Time***, ***Stance Time*** and ***Swing Time***. These parameters are typically measured using an instrumented treadmill, force plates or a motion capture system as the gold standard [16]. These four parameters are estimated by WalkEar.

Spatial gait parameters relate to the body movements of an individual while walking. An important spatial parameter, the ***Vertical Displacement*** (VD) refers to the up-and-down movement of the body's center of mass during walking.

In summary, spatio-temporal gait parameters measured in WalkEar are commonly used parameters that aid in the understanding of changes in gait timing and fitness tracking in daily life. For example, a study showed how cadence is linked to the metabolic zone and intensity [17]. Spatio-temporal parameters also play a valuable role in clinical settings. For example, variability in parameters like cadence, swing time, and stance time is critical for diagnosing Parkinson's Disease [18], requiring step-to-step assessment as is performed in WalkEar.

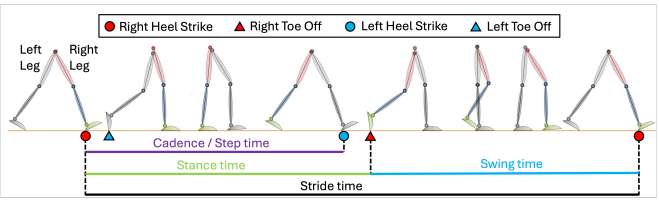
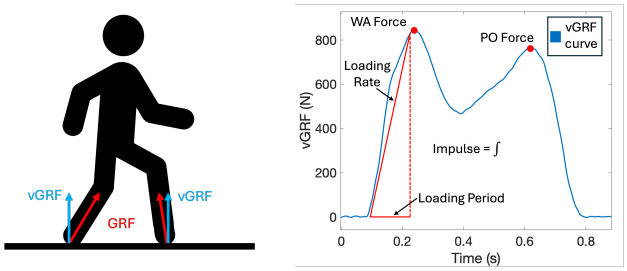


Fig. 1: Illustration of a gait cycle and corresponding temporal parameters. Adapted from [19] with Creative Commons licence.



(a) (b)

Fig. 2: (a) Visualisation of the GRF and vGRF while walking. (b) A typical walking vGRF curve annotated with the scalar kinetic parameters assessed by WalkEar.

B. Kinetic gait parameters

Kinetic gait parameters are derived from the interaction between the foot and the ground. When the foot makes contact with the ground, a force called the ground reaction force (GRF) is exerted by the ground onto the centre of pressure of the foot. The vertical ground reaction force (vGRF) represents the greatest force exerted by the ground on the body, as shown in Figure 2a. It provides valuable insight into weight transfer through the lower limbs, load bearing in the body and gait style. It is also linked to various pathologies [20]. Within laboratory settings, the gold standard method of vGRF measurement is via the use of floor-embedded force plates, or force plates integrated into a treadmill. By measuring the vGRF over time, we can obtain a vGRF curve, as illustrated in Figure 2b. This curve can be labeled with scalar characteristics, particularly the two peaks, the ***Weight acceptance force*** and ***Push off forces***, the initial slope called the ***Loading Rate*** and the area under the curve, known as the ***Impulse***.

C. Asymmetry

Each gait parameter listed above is associated with a left-right asymmetry. Asymmetry is important for identifying gait abnormalities and monitoring changes in symmetry over time, such as the onset of disease, during rehabilitation, or when undergoing treatment [6].

We report symmetry using a symmetry index (SI) [21], [22]. In this work SI is given as a percentage to allow easier interpretation and can be calculated as follows : $SI = \frac{X_{left} - X_{right}}{X_{reference}} \cdot 100\%$, where $X_{reference}$ is the average value of X_{left} and X_{right}. When calculating the symmetry index, the choice of reference depends on what is being assessed. For example, in the case of injury, the healthy side is often used as the reference.

III. System Design

The overview of WalkEar is shown in Figure 3. It starts by taking two-channel IMU data as input. A series of preprocessing steps are applied, followed by the estimation of temporal parameters. The detected gait events are then utilized for step-by-step estimation of gait parameters. Finally, based on

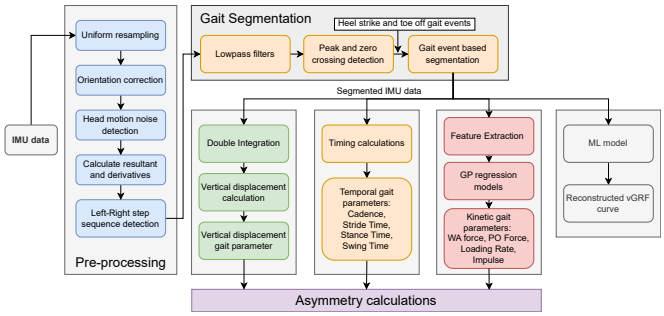


Fig. 3: Flowchart of WalkEar system and gait parameters

the estimated spatiotemporal and kinetic parameters, WalkEar performs an asymmetry analysis.

A. Pre-processing

1) Coordinate system: WalkEar utilizes IMUs on both sides. The coordinate system of each side defined in WalkEar, denoted as $Coor_g$, as illustrated in Figure 4, consists of three axes: Vertical (VT), Anterior-Posterior (AP), and Medio-Lateral (ML). After capturing the IMU data, it is resampled from a non-uniform rate to a uniform sampling rate of 150 Hz.

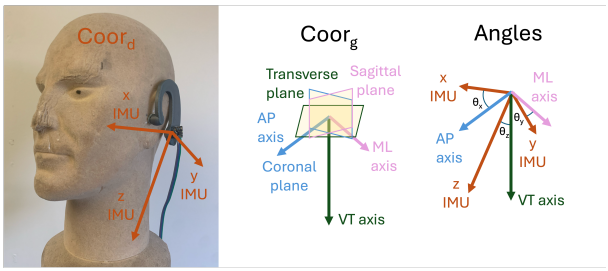


Fig. 4: Depiction of the coordinate systems $Coord_d$, shown on a head with a 3D printed earable. $Coord_g$ shown with the anatomical planes. The angles θ_x, θ_y and θ_z shown between $Coord_d$ and $Coord_g$.

2) Opportunistic calibration of earphone orientation: When a user wears the earphones, the device coordinate system, denoted as $Coor_d$, may differ from $Coor_g$, due to user movements or the wearing orientation. Since the IMU data is collected in $Coor_d$, we need to project it into $Coor_g$. In WalkEar, we conduct opportunistic calibration of the earphone orientation to compensate for coordinate misalignment utilising the measured components of gravity in each IMU axis. This process occurs when the user starts to walk and during any periods of being stationary while walking. The workflow is:

- First, we identify the stationary period where the resultant acceleration is near gravity, i.e., $G = 0.981$ m/s, where the resultant acceleration is $a_d^t = \sqrt{(a_x^t)^2 + (a_y^t)^2 + (a_z^t)^2}$, and a_x^t, a_y^t and a_z^t are the acceleration on three axes of $Coor_d$ at time t respectively. When the user is stationary, the mean value of acceleration during a small window in the AP and ML directions of $Coor_g$ should be 0, and the mean value in the VT direction of $Coor_g$ should be close to gravity, i.e., G.
- We then calculate for each axis i:

$$\theta_i = arccos\left(\frac{mean(a_i^t)}{G}\right) \quad (1)$$

where θ_i and a_i^t are repeated for the rotations angles between x, y, and z axis of $Coor_d$ and VT of $Coor_g$ respectively as shown in Figure 4.

- Finally, we transform the rotation angles $(\theta_x, \theta_y, \theta_z)$ into a quaternion for coordinate system transformation between $Coor_d$ and $Coor_g$ separately for the left and right ears.

Figure 5 shows an example of the acceleration data before and after the transformation while stationary. It demonstrates how this transformation aligns the left and right devices, which initially differ due to variations in how they are worn by the user on each side, into the $Coor_g$ coordinate system.

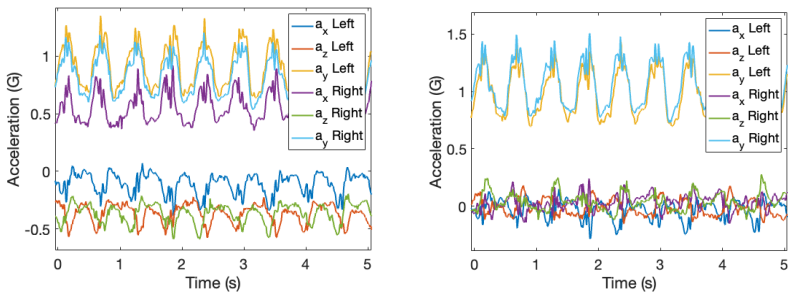
3) Head motion artifact removal: When the user is walking, unexpected head motions can affect the IMU data, degrading the performance of gait monitoring. We make a key observation that when there are head motions, the gyroscope data is significantly stronger than when the user is walking, due to the much smaller radius of rotation for the head motions. Figure 6 shows the regions of gyroscope data above the threshold, chosen as 3 rad/s, during a period of walking. These sections with high amplitude gyroscope signals correspond to head motions. The gait parameters from these gait half cycles (from one HS to the next HS) are discarded and an average of the surrounding gait cycle parameters are instead used. To account for the significant yaw drift over time in the IMU gyroscope, a highpass filter is used to remove the baseline wander so head rotations are accurately found. These filtered gyroscope signals are presented in Figure 6 where short-term variations can be seen but no long-term drifts.

4) Left-Right step detection: Since WalkEar utilises IMU data from both sides, it can provide left-and-right asymmetry analysis and detailed kinetic gait parameter analysis for each side. To achieve this, we first subtract the right-side data (from the right earbud) from the left-side data, then apply a low-pass filter with a cutoff frequency of 5 Hz to filter out the high-frequency noise. As illustrated in Figure 7, when the user leans towards the left, indicating a left step (i.e., left HS), the filtered data exhibits two large positive peaks. When the user leans towards the right, indicating a right step (i.e., right HS), the filtered data shows two large negative troughs. Therefore, we conduct peak (trough) detection on the filtered data to recognize each left (right) step. This goes to a sequence detector which looks for two positive peaks, and then two negative troughs, ignoring a trough or peak in between respectively.

B. Temporal gait parameter estimation

To measure the temporal gait parameters from pre-processed IMU data, it is key to detect the timing of two gait events on both sides: the HS and the TO, as illustrated in Figure 1. We demonstrate the proposed HS and TO detection method on one side, but the same process applies to both sides.

HS event detection. Since the HS occurs when the heel contacts the ground, it creates a significant spike in accelerometer



(a) (b)

Fig. 5: Acceleration signals from the left and right IMU (a) before and (b) after rotations.

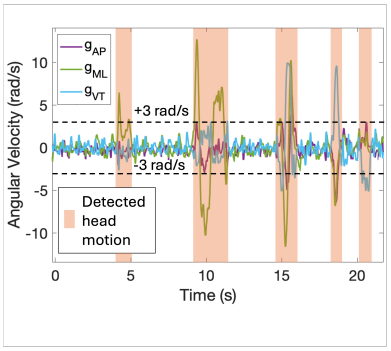


Fig. 6: Example of head motion removal.

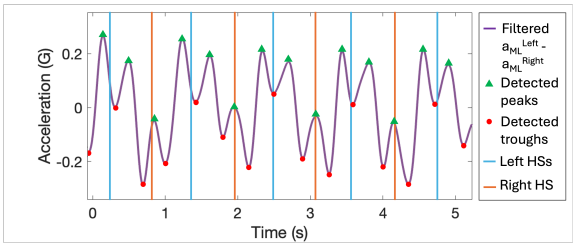


Fig. 7: Filtered (a_{ml}^{left} - a_{ml}^{right}) signal with detected peaks and troughs, showing an alternating pattern corresponding to left and right steps.

data. We thus calculate the resultant acceleration, denoted as $a_g^t = \sqrt{(a_{ap}^t)^2 + (a_{ml}^t)^2 + (a_{vt}^t)^2}$, where a_{ap}^t, a_{ml}^t and a_{vt}^t are the acceleration on three axes at time t. Then, a_g^t is differentiated to obtain a jerk signal. Peaks in the jerk signal correspond to HS events, as these events are characterized by sharp changes in acceleration when the heel impacts the ground (illustrated in Figure 8). We observed that peaks in the acceleration itself lagged behind the actual heel strike event, resulting in a weaker correlation with cadence and stride time when used as the feature for identifying HS.

TO event detection. The TO occurs when the toes leave the ground, causing the foot lose contact with the ground. TO detection is difficult as it is very close to the opposing legs' HS which is a much stronger signal. To detect TO events, we analyze the acceleration along the VT axis (a_{vt}^t) because changes in gravitational acceleration are indicative of the foot's movement away from the ground. Specifically, we first calculate the first-order difference of a_{vt}^t (denoted as a'_{vt}), to calculate the jerk signal. After that, a Butterworth low-pass filter is applied, with a 2 Hz cutoff frequency, to aggressively smooth the resulting data (denoted as $\sim a'_{vt}$). Our algorithm then identifies troughs and zero crossings in the smoothed data following the most recent HS to identify the TO. In the smoothed data, troughs indicate a locally large change of acceleration, which is typical when the foot transitions from being in contact with the ground to lifting off. Thus, the troughs are indicative of a TO point. The zero crossing is used to locate the correct trough to use, as these peaks represent stationarity in the TO acceleration where the foot is finished lifting upwards. This process is shown in Figure 9 where the TO events given in red lines are shown to align with the minimum points followed immediately by a zero crossing in the $\sim a'_{vt}$ signal.

Parameter estimation. After detecting the HS and TO events on each foot, the temporal gait parameters listed in Section II-A can then be calculated using the differences in timings of gait events shown in Figure 1. These estimations rely only on the most recent gait event timings, rather than a summation of prior timings, thus making WalkEar robust to IMU sensor drift and error accumulation.

C. Spatial gait parameter estimation

The estimation of vertical displacement during one step involves a multi-step process by capturing the vertical oscillations of the center of mass. The algorithm utilises numerical integration to move from acceleration, to velocity, and displacement. Between each integration step, a high-pass filter is used to remove the drift of the sensor as well as the offsets in integration. Finally, the data is segmented using the HS gait events detected previously. The difference between maximum and minimum displacement is calculated and given as the Vertical Displacement parameter. Figure 10 illustrates the left, right and average estimates of vertical displacement from WalkEar against the ground truth. It can be seen that the purple "average" curve tracks the ground truth better than either the left or right earables. This indicates that using both earables gives a better estimate of movement of the center of mass than a single earable. This estimation similarly happens over a single gait cycle, lasting around half a second, reducing any impact of sensor drift on the measurements.

D. Kinetic gait parameter estimation

The head does not replicate the characteristics of the acceleration or force measured at the center of the foot. This distinction is evident in Figure 9, which displays both earable acceleration (yellow) and ground reaction force (blue and orange). The significant differences between these signals show that a simple proportional relationship, such as one based on mass, cannot link them effectively, motivating the use of regression techniques.

Specifically, this can be seen in Figure 9 showing both earable acceleration (yellow) and the ground reaction force (blue and orange). These signals are shown to be very different

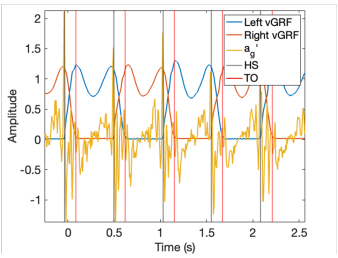


Fig. 8: Derivative of a_g, or the jerk signal, with labelled heel strike events, showing peaks corresponding consecutive to heel strikes.

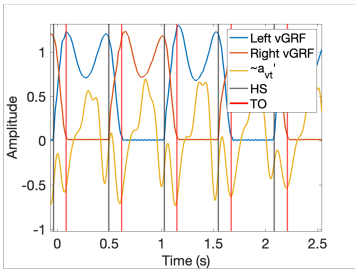


Fig. 9: Filtered a_{vt} ($\sim a'_{vt}$) derivative signal with labelled toe off events, showing troughs corresponding to consecutive toe off events.

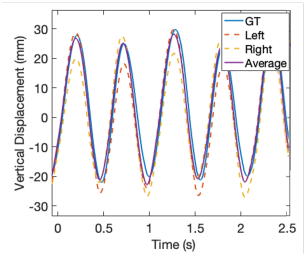


Fig. 10: Vertical displacement estimated by the left and right earables, and the average signal between them, compared to the GT.

from each other, showing that a constant of proportion like mass will not link the signals together, hence motivating the use of regression techniques.

1) WA Force, PO Force, LR, and Impulse estimation:

Segmentation. We utilize the resultant acceleration, a_g^t, to estimate kinetic gait parameters because its amplitude is directly related to the force exerted. The acceleration is segmented for regression analysis. Specifically, for the estimation of PO Force and Impulse, the data is segmented for each stance phase. For WA Force and LR estimation, we select the period from the HS to 25% of the stance period (usually close to 0.15 s). This is because the WA Force and LR are present at the start of the stance period as shown in Figure 2b, compared to the PO Force which occurs near the end, and the Impulse which is derived over the whole stance period.

Feature extraction. The features extracted from each segment are chosen for their ability to capture the dynamics needed for kinetic gait analysis, validated with an F-Test. These features encompass both time and frequency domain metrics, each offering unique insights into gait mechanics. The time domain features are maximum amplitude, mean and RMS values, shape factor and crest, and clearance and impulse factors, which directly measure the magnitudes and variability of forces exerted during gait events. The frequency domain features are Mean Frequency, Peak Amplitude and location, and Band Power.

Regression model. To estimate the scalar kinetic gait parameters, we employ an exponential Gaussian Process (GP) regression model [23] using the extracted features to predict the four kinetic gait parameters: WA Force, PO Force, LR, and Impulse. Each parameter is estimated using a dedicated GP model, trained specifically for that parameter. The models are trained using ground truth forces normalised by the user's body weight. To personalise the model, separate models are trained and evaluated using ground truth and inputs in Newtons (i.e. unnormalised forces).

2) vGRF curve reconstruction: Histogram-based Gradient Boosting Regression models are used to predict the vGRF curve. This approach is selected for its efficiency in handling sequence prediction tasks while maintaining relatively low computational overhead. The input is the same as the scalar parameter models segmented as described previously. The training data ground truth uses force data that is normalised by the participants' body weight so that predictions are made in normalised forces with the unit body weight (BW). These can be scaled to forces in Newtons if the user's weight is known, however, it is common in the literature to estimate the vGRF profile only in the BW unit [24]. We trained two models for WalkEar at different levels of granularity, the first model uses 20 samples and the second uses 100 samples, giving a finer resolution but increased computational overhead. It is shown in Section V-I that the 20 sample model can run in real-time on a phone.

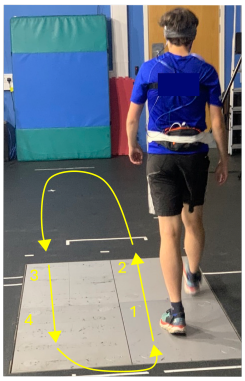
IV. Data Collection and Protocol

We collected data from 13 participants while walking. The cohort included 9 males and 4 females amounting to 130 minutes of treadmill data and 39 minutes of force plate validation data, resulting in over 18,000 step samples. To the best of our knowledge, no similar public dataset exists to evaluate our system. The participants had a mean age of 29 mean weight of 73.0 kg mean height of 1.78 m mean BMI of 21.5and mean foot length (calculated from shoe size) of 27.1 cm. Each participant was instructed to wear running shoes for the experiments. This study was approved by the University of Cambridge Department of Computer Science and Technology Ethics Committee under application number 2134. All participants in the study were healthy, however, variance in gait parameters was created by a range of walking speeds and heights of the participants, we show the mean and standard deviation of each gait parameter in Table I. The participants wore a 3D-printed earbud with an IMU mounted to the earpiece. We used the MPU6050 [25] IMU which has an inbuilt accelerometer and gyroscope. The accelerometer range was set to ± 2 G which encompassed the full range of walking signal, and the gyroscope to $\pm 250\,^{\circ}$/s. The data collection protocol consisted of a 5-minute walk at 3 kph followed by 5-minute walk at 5 kph on a gold standard [26] Bertec instrumented treadmill. Then participants performed the in-the-wild test, completing 3 minutes of free walking at an uncontrolled speed in a 15x5 m sized room with floor-embedded force plates. On the treadmill, the participants were kept at a constant 0% incline to mimic the dynamics of real walking [27]. Ground truth data was collected at a sampling frequency of 1000 Hz and pre-processed with a lowpass filter with a cutoff frequency of 50 Hz. The participants were not instructed to keep their head stable and were allowed to look around the

TABLE I: Mean and standard deviation (STD) values from the collected dataset for each gait parameter.

Parameter	Cadence	Stride	Stance	Swing	WA Force	PO Force	LR	Impulse	VD
Mean Value	111 step/m	1.08 s	0.680 s	0.418 s	849 N	824 N	8860 N/s	403 Ns	0.73 cm
STD	12 step/m	0.14 s	0.026 s	0.038 s	166 N	148 N	2200 N/s	88 Ns	0.17 cm

(a) A participant standing on the instrumented treadmill. Note: during walking, both belts are used otherwise the left and right steps intersect.



(b) The force plates annotated with the user's path. The force plates are labeled in the order the user steps onto the plates.

Fig. 11: Annotated photographs showing the experimental setup used to validate WalkEar.

lab, although this was not specifically tested for. Additionally, ground truth for vertical displacement was collected with a Qualisys motion capture system using markers attached to the participant. The experimental setups are shown in Figure 11.

V. Evaluation

A. Gait parameter estimation performance

1) Timing: The results for the temporal gait parameters (*i.e.,* cadence, stride time, stance time, and swing time) are shown in Table II. We detect cadence with excellent performance, a MAPE of 2.06%, amounting to an error of 2 steps per minute within an average cadence per person of 111 steps/min in our dataset. We detect stride time with a 1% error. Stance time and swing time are computed with low errors of 3.1% and 5.1% respectively. We observe that stance time error is higher than that of swing time, on account of the complexity of accurately detecting the toe off points. Although the heel strike creates large spikes in the accelerometer signal, the toe off motion is more subtle and needs to be identified using intelligent processing. Despite this added complexity, we achieve results that are competitive with literature (as shown in Section V-B).

2) Vertical Displacement: The estimation results for vertical displacement (VD) are provided in Table II. We achieve a 1.93mm (or 2.72%) error in estimating the vertical displacement compared to the gold-standard vision system. To fully evaluate system performance, we also estimated vertical displacement without earphone orientation calibration. Without calibration, the error in vertical displacement is 12.2%, highlighting the importance of the orientation correction.

3) Kinetic: Overall results for each kinetic parameter examined are provided in Table II. These are validated using a leave-one-subject-out (LOSO) validation scheme to simulate the most likely application scenarios where no user-specific training data is available, due to the expense and complexity of label collection using high-end treadmills or force plates. The presented results are averaged over all participants. Overall, we achieve very low errors of less than 4% for all kinetic parameters. Additionally, as shown in Table IV, we are able to predict the kinetic parameters with good agreement with the ground truth. This shows that we are able to predict the variance in the kinetic parameters well and with little error.

4) Gait Asymmetry: This section reports on the accuracy of the asymmetry of the gait parameters estimated by WalkEar. The asymmetry values are reported as a symmetry index (where the reference value is taken as the left-right mean). Table III shows the absolute mean GT asymmetry, the absolute mean estimated asymmetry, and the MAE between them. We report the asymmetry in overall timing, asymmetry in each kinetic parameter, and the overall kinetic asymmetry. We show that our system is sensitive enough to detect the changes in parameters between the two sides of the body with very low errors for all asymmetries and always estimating the correct left-right direction. Thus, our system is applicable to asymmetry monitoring even when the extent of the asymmetry is small.

B. Comparison with related works

Table IV shows a comparison of the performance of WalkEar with the reported performance of other earable gait parameter papers discussed in Section VI. We implement the algorithm of the best performing related work on temporal gait parameters, EarGait [15], in our dataset, as well as using their reported results. For the e-AR sensor we take the best result from the three works using the sensor.

From Table IV, it is evident that WalkEar outperforms the related work on the timing-related parameters. Specifically, on the more difficult task of stance/swing time estimation, the performance of WalkEar is far superior to the best found in the literature (with an improvement of 44% and 41% for stance and swing times respectively over the best results in Table IV). WalkEar is also comparable to EarGait [15] in cadence and stride time, which only require localisation of the heel strike. Similarly, when the EarGait algorithm was applied to our dataset, WalkEar demonstrates comparable performance in cadence estimation and superior performance in estimating other parameters, with EarGait showing marginally worse results than the reported results in the paper [15].

Kinetic parameters have only been studied from the ear by Attalah et al. [29] using the e-AR earable sensor. To compare, we calculate the R^2 score, which is a measure

TABLE II: Overall results for all gait parameters evaluated across the whole dataset.

Metric	Cadence step/min	Stride Time s	Stance Time s	Swing Time s	VD mm	WA Force N	PO Force N	Loading Rate N/s	Impulse Ns
MAE	2.47	0.0114	0.0216	0.0215	1.93	17.5	8.59	328	2.26
SDE	4.55	0.0214	0.0313	0.0314	2.51	23.2	11.3	491	3.59
MAPE	2.06%	1.02%	3.10%	5.14%	2.72%	2.12%	1.34%	3.74%	0.73%
ME	-0.21	-0.00014	-0.00694	0.0067	0.94	2.69	-0.087	-105	-0.24

TABLE III: Overall asymmetry index.

Parameter	Symmetry Index (%)		
	GT	Earable	MAE
Timing	1.76	2.78	1.38
WA Force	4.72	4.17	0.53
PO Force	3.45	2.95	0.66
LR	22.1	19.29	3.11
Impulse	2.01	1.87	0.19
Overall Kinetic	4.41	3.84	0.60

TABLE IV: Baseline comparison.

Parameter	Ours	e-AR [28], [29], [14]	EarGait [15] Reported	EarGait [15] Our Dataset
Cadence MAE	2.47	8.88	**1.85**	1.93
Stride time MAE	**0.011**	0.028	0.012	0.015
Stance Time MAE	**0.022**	0.038	0.076	0.064
Swing Time MAE	**0.022**	0.036	0.078	0.067
WA force (R^2)	**0.69**	0.35	-	-
PO Force (R^2)	**0.65**	0.36	-	-
Loading Rate (R^2)	**0.73**	-	-	-
Impulse (R^2)	**0.69**	0.26	-	-

of the goodness of fit of a regression model. To obtain this result, we performed 5-fold cross-validation on data from all users per kinetic parameter to mirror the approach used by Attalah et al. [29] which, importantly, did not assess their system with held-out users. It is clear that WalkEar significantly outperforms this work with far superior R^2 scores for each parameter, as well as addressing new parameters. In summary, WalkEar outperforms earable literature on both temporal and kinetic parameters, including when the best performing related work is implemented on our dataset. In addition to improved performance, WalkEar intelligently leverages features from two earbuds, while the earable related works use only a single earable.

C. vGRF curve reconstruction

The evaluation of the vGRF curve estimation was performed with 3 validation strategies. A Leave-one-subject-out (LOSO) cross-validation, as was done with the prior kinetic parameters. As well as both 5-fold cross-validation across a mixture of user data and per-user validation were performed to compare with related work that doesn't hold out one user. We report error as Normalised Root Mean Square Error (NRMSE) on this task to compare with related work. We trained two different models for vGRF reconstruction at different granularities, one with 20 samples per step (referred to as 20), and one with 100 samples (referred to as 100). These two versions are provided to balance computational efficiency and granularity. We provide the results of vGRF reconstruction in Table V and also compare our results to those obtained by Jiang et. al. [24], who estimated the vGRF curve from IMUs placed on the lower leg, feet and trunk.

TABLE V: Comparison of vGRF curve reconstruction.

Validation Method	NRMSE			Correlation		
	Ours (20)	Ours (100)	[24]	Ours (20)	Ours (100)	[24]
LOSO	5.25	5.12	-	0.985	0.983	-
5 Fold CV	**2.23**	**2.30**	7.15	**0.995**	**0.996**	0.97
Individual	8.67	8.67	**1.7**	0.958	0.958	**1.00**

From the table, it is clear that we achieve good performance across each validation strategy with very high correlation between the predictions and the ground truth curves. Interestingly, the 5-fold cross validation and leave-one-subject-out perform better than the individual model. This is since the quantity of data for training is much smaller for an individual model, but this performance is expected to increase with more data per user. We also see that the 20 sample and 100 sample models achieve similar results. However, as illustrated in Figure 12, the 100 sample curve provides more detail than the 20 sample model. There is thus a trade off between greater detail but larger model size (88 MB) and inference time in the 100 sample model compared to the 20 sample model with less detail but smaller model size (18 MB). We also see from the table that WalkEar outperforms the existing hip and foot based wearables from the literature [24] for 5-fold cross validation, and that our leave-one-subject-out performance is better than the 5 fold cross validation from the literature. This shows the power of our system to generalise to new users.

When evaluated against the GP models in estimating the four scalar kinetic parameters this method performs worse. Our evaluation shows a 4% increase in error across the four parameters on average from the vGRF curve vs the GP models. This is thought to be due to the downsampling of ground truth vGRF curve for model training. This affects precise peak heights, integration and differentiation for WA Force, PO force, Impulse and LR respectively. Therefore, we use scalar kinetic parameter estimation models for better performance and lower computational overhead. However, the vGRF curve reconstruction itself provides much finer detail on the vGRF profile, such as peak and trough locations in time and inflections during the loading period.

D. Individualised performance

In this section, we examine the performance per user. Figure 13 provides results for the average MAPE of the

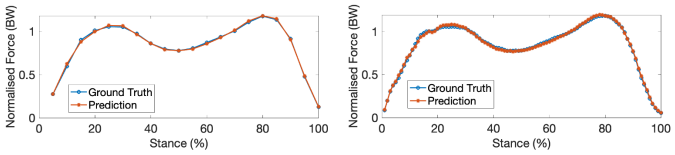


Fig. 12: Sample result for the vGRF reconstruction from the test data using the Leave-one-subject-out validation method for the (a) 20 sample model and (b) 100 sample model.

temporal (blue) and kinetic (orange) parameters for each participant. The error in kinetic parameters is consistently low amongst participants. However, the temporal parameters have more variation with Participant 7 having the largest error of 7.1% and Participant 1 having the lowest error of 0.87%. Participant 7 has a larger error for temporal parameters due to a loosely fitting IMU which dampens the vibrations of the gait cycle leading to a worse estimation of the TO event. In addition, to assess the impact of body type upon the errors of WalkEar, we performed a t-test of the individual MAPE against each participant's height, weight BMI and foot length. The results of this analysis showed no correlation at the 5% significance level, although the sample size of this experiment was limited to 13 participants. These results may suggest that individual errors in WalkEar performance are linked to the walking style of the individual rather than the body type.

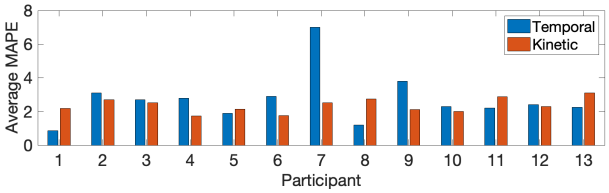


Fig. 13: Comparison of average MAPE for averaged temporal and kinetic parameters over each participant.

E. In-the-wild experiments

This section details two experiments ran in less controlled settings to evaluate the robustness of WalkEar in in-the-wild conditions. Accurate GRF ground truth data requires force plates, which are typically embedded in treadmills, as used to collected the majority of the data in this study. Collecting precise GRF data without force plates is not feasible, limiting options for in-the-wild experiments. For instance, insoles equipped with force-sensitive resistors are unsuitable for providing ground truth data. These devices have been shown to produce errors in vGRF predictions comparable to the reported performance of WalkEar [30], making them ineffective for evaluating WalkEar.

1) Free walking validation: To assess system performance under less controlled settings, we tested participants' free walking on force plates on the floor where the participants freely chose and varied their walking speeds in the experiments as illustrated in Section IV. The walking speeds had an average

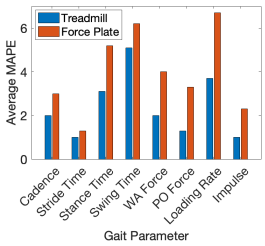


Fig. 14: Comparison of the errors during free walking and on a treadmill.

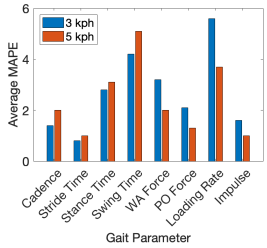


Fig. 15: Error comparison of all gait parameters with different treadmill speeds.

standard deviation of 0.83 kph compared to 0.07 kph on a treadmill, showing a significant increase in gait speed variability in the free walking experiment. Participants walked in laps around a room 5x15 m in size over four floor-embedded force plates (an illustration of this is found in Figure 11b). The results for all parameters are presented in Figure 14. From the figure, a performance degradation for both temporal and kinetic parameters can be seen, with a higher percentage change for the kinetic parameters. This is thought to be due to the kinetic parameters being estimated from ML models thus experiencing a distribution shift when being tested on force plate data when trained on treadmill data, compared to the temporal parameters being estimated with signal processing methods. For this experiment, ground truth on vertical displacement was not collected so it is not presented here. Overall, the performance during free walking shows only a slight increase in **MAPE** of **1.6%** for temporal parameters and **2.1%** for kinetic parameters compared to controlled treadmill experiments. We also perform an analysis of the impact of walking speeds during free walking as presented in Section V-F.

2) Stop-and-go scenario validation: WalkEar was tested with a stop-and-go scenario, where participants transitioned from standing still to walking. To acquire reliable ground truth information for evaluation, this experiment was conducted on the Bertec treadmill, as the force plates would not allow enough consecutive ground truth steps to reach the target speed. Participants started walking on a treadmill from stationary and gradually increased to their desired speed, typically taking 5-8 steps to reach the target. Overall, the experiment demonstrated an increase in **MAPE** of **7.3%** for temporal parameters and **5.6%** for kinetic parameters. This indicates a higher error rate compared to the treadmill experiments, however, the errors mainly occurred during the initial steps and settled quickly thereafter, still marking an improvement over related works.

F. Performance impact from differing walking speeds

Treadmill speed variations: The user's walking speed affects the strength of the vibrations propagated to the head during walking. Figure 15 compares system performance when walking at two different speeds. It can be seen in the figure that the error is smaller under slower speeds for timing parameters, but larger under slower speeds for kinetic parameters. For timing parameters, this smaller percentage error is mainly due

to the longer time taken for each gait cycle. For the kinetic parameters, weaker foot strikes of slower walking result in weaker vibrations being captured by the IMU making the estimation of kinetic parameters more difficult.

Free walking speed variations: In addition, we performed an experiment to test WalkEar's performance under varying speeds in the free walking experiment. Participants were divided into two groups based on their average walking speed: above and below the median speed of 4.3 kph. WalkEar's performance was then evaluated separately for each group. The results displayed a pattern similar to that observed in the treadmill speed analysis. In the faster group the **MAPE was 4.17%** for temporal parameters and **3.76%** for kinetic parameters. In the slower group, the MAPE was **3.82%** for temporal parameters and **4.25%** for kinetic parameters. Overall, WalkEar performs reliably under various speeds in free walking.

G. Changes in footwear

All the data used for WalkEar was collected with participants wearing running shoes. However, to show the impact on WalkEar to the user's footwear choice, we compare the previously presented results to an experiment where a participant wore flat style shoes for a five minute walking experiment. The results for this experiment are given as the MAPE values for each gait parameter subset. The **MAPE** for spatiotemporal parameters for flat shoes is **4.06%** compared to **3.94%** for running shoes showing only a marginal increase for this participant. The **MAPE** for kinetic parameters was **3.34%** for flat shoes and **2.45%** for running shoes. While the increase in MAPE for kinetic parameters was higher than that for spatiotemporal parameters, it was still within one standard deviation of the overall MAPE across the dataset. We expect that this increase is due to all the training data for kinetic parameters being on running shoes rather than a mixture of footwear.

H. Benchmark evaluation

To fully understand the performance of WalkEar, we evaluate the performance under different settings.

1) Impact of both channels: This section assesses the impact of gait parameter estimation when wearing two earables, versus just one. We present the results of this analysis in Figure 16, which provides the performance using only the left, right and fusion of both earbuds for each parameter. It is evident from the figure that the fusion of both earbuds results in significant performance gains. However, it is also evident that each channel can be used in isolation with reasonable performance.

2) Impact of IMU drift: To show that the WalkEar performance does not degrade due to IMU sensor drift, we compare the performance during the first and last minute of a five minute long walking experiment. This was repeated and averaged over 6 different participants and the averaged results are presented in Table VI. From the Table it can be seen that the MAPE does not significantly change between the start and end of the recordings, showing the algorithm does not accumulate error over time. Specifically, this is also true for the spatial parameters that uses the integral of the acceleration signals.

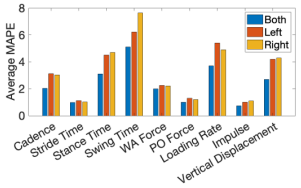


Fig. 16: Results for overall MAPE using both, left or right IMU channels for each gait parameter.

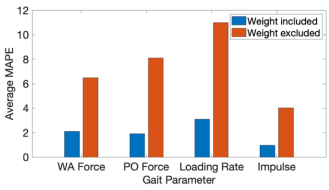


Fig. 17: Overall error of kinetic gait parameter estimation with and without personalisation.

TABLE VI: MAPE and standard deviation of MAPE (given as MAPE±SD) for the first and final minutes of a five minute walking experiment, grouped by subset of gait parameter.

Gait parameter subset	Temporal	Spatial	Kinetic
First Minute	2.91%±3.78%	2.64%±2.93%	1.97%±3.02%
Final Minute	2.79%±3.45%	2.83%±3.12%	1.93%±2.83%

3) Sampling rate reduction: This experiment assesses the impact of downsampling on estimation errors. We see that, as expected, as the sampling rate is decreased, performance slightly decreases. However, this decrease does not significantly degrade performance, with a maximum error amongst all the parameters at a 20 Hz sampling rate of 9% for swing time. This is important since it means that our system can be run on commercial earables which might impose lower sampling rates due to battery life constraints, such as the AirPod Pro with a maximum IMU sampling rate of 25 Hz [31].

TABLE VII: Impact of downsampling the IMU signal.

Sample rate	150 Hz (original)	100 Hz	50 Hz	20 Hz
Cadence	2.06%	2.12%	2.46%	3.84%
Stride time	1.02%	1.06%	1.11%	1.75%
Stance Time	3.10%	4.35%	5.52%	7.59%
Swing Time	5.14%	6.34%	7.82%	9.27%
WA force	2.12%	2.91%	3.14%	4.80%
PO Force	1.34%	1.95%	3.14%	4.76%
Loading Rate	3.74%	4.85%	5.43%	6.51%
Impulse	0.73%	1.42%	1.99%	2.89%

4) Personalisation: When determining the kinetic parameters (as reported in Table II), we apply personalisation using the user's weight. When training our model, we weight normalise the kinetic parameters and then scale the outputs by the user's weight, resulting in a final prediction of force in Newtons (N), Loading Rate in Newtons per second (N/s) or impulse in Newton Seconds (Ns). In this section, we assess the impact of personalisation by providing error when training with weight normalisation as has been presented in earlier sections versus that when training without in when the system knows the user's weight Figure 17. Significant performance improvements are obtained by the model knowing the user's weight, showing the power of our personalisation technique. User weight is easily acquired from the users themselves and does not dramatically change over time so it is an acceptable input for WalkEar.

I. System Performance

While WalkEar analysis can be done in the cloud, we believe it would be preferable in some situations (i.e. for privacy preservation) to perform computation on a personal device, e.g., a phone. Additionally, some on-device applications such as user identification [32] and biometrics [33] use live gait information. Therefore, WalkEar aims to provide a privacy-conscious, on-device solution, ensuring that the algorithms remain lightweight and responsive without compromising privacy.

TABLE VIII: Power consumption and latency.

Algorithm stage	Battery consumption in 30 minutes (%)	Latency (ms)
Preprocessing and Timing	2%	0.006
Kinetic	3%	65.7
vGRF (20 samples)	10%	328.5
Spatial	2%	0.012
Overall	12%	394.22

This section presents the power consumption and latency of predictions for WalkEar when executed on an iPhone 15 Pro (3290 mAh battery). We created a phone application which runs our algorithms, and present the resultant system performance in Table VIII. We report the latency for each group of parameters per step. The average cadence in our dataset is 111 step/min, amounting to approximately 0.5 step/second. Our system can operate in real-time since our overall system latency is 0.39 s (less than 0.5 step/second). The 20-sample vGRF curve achieves this real-time processing, while the 100-sample curve, with a 1.6 s latency, is suitable for offline analysis with higher resolution requirements. Communication latency between the earbud and phone is negligible, with latency dominated by processing time. Battery consumption was measured by running the application continuously for 30 minutes. The application consumed 12% battery, primarily due to the high latency of the vGRF sequence-to-sequence model. However, if kinetic parameter estimation without curve reconstruction is sufficient for the user's purpose, the overall battery consumption will be significantly decreased. For context, idling with the screen on for 30 minutes consumes 1% of the battery. Running timing and spatial parameters increases battery usage by an additional 1%, while kinetic parameter estimation adds 2%. The GP models for kinetic parameter estimation have a size of 700 kB, while the vGRF reconstruction models are 18 MB (for 20 samples) and 88 MB (for 100 samples). The lightweight nature of our models allows WalkEar to run feasibly within the memory constraints of a typical smartphone.

VI. Related Work

A. Wearable-based Gait Analysis

Gait analysis has been studied using various wearable devices such as smartwatches [11], belt-attached pods [9], smartphones [34], [35], and smart insoles [7], [36]. These devices estimate parameters such as step length, swing time, and stance time. Some studies have explored vGRF curve reconstruction using multiple inertial sensors [24], [37]. However, these methods often require multiple sensors or custom devices, limiting their practicality and user-adherence for daily use.

B. Earable-based Gait Analysis

Commodity earables, such as Apple AirPods Pro [38] and high-end hearing aids [39], are equipped with IMU sensors. Previous earable-based gait studies have focused on activity recognition [40], step counting [41], and gait pose classification [42]. Research on gait parameters using earables includes work on the e-AR sensor for temporal parameter estimation [43]. Jarchi et al. [28] developed an algorithm using singular-spectrum analysis (SSA) to detect heel strike and toe-off events. Subsequent studies [14], [15] further refined this algorithm. Additionally, asymmetry analysis was investigated from this algorithm [13]. For kinetic parameters, Attalah et al. [29] achieved weak correlations using peak amplitude features from earable IMU signals. These works are used for comparison to WalkEar in Section V-B. The existing earable-based studies typically focus on measuring isolated subsets of walking gait parameters for specific applications. In contrast, WalkEar is the first earable-based system that supports application-agnostic estimation of a comprehensive suite of spatial-temporal and kinetic parameters, offering detailed step-to-step analysis. Additionally, WalkEar estimates the full vGRF curve from earable IMUs, showing comparable results to studies using lower limb and trunk-mounted IMU sensors [24].

VII. Conclusions

We have presented WalkEar, an application-agnostic walking gait monitoring system using IMU data from earables. WalkEar achieves comparable or superior performance to existing literature across all tasks, providing continuous and detailed analysis of spatio-temporal, kinetic parameters, and accurate asymmetry analysis. Notably, we demonstrate for the first time that vGRF can be estimated from the ear with comparable RMSE to sensors on legs and shoes. Future work should include validations on uneven surfaces and longitudinal experiments to explore WalkEar's effectiveness in observing gait patterns and deviations over time. While WalkEar is not tested on gait signatures from participants with pathological gait, as per the analysis conducted in Section 4, it can measure a varied range of gait parameters as it operates on a step to step basis with no averaging and uses gait events with distinct features in the IMU signals to segment the gait cycle. This opens the door to the application of monitoring pathological gait. Additionally, our dataset contains different speed levels as well as different participants, some with some natural imbalance in gait, creating a large variance in the dataset that WalkEar correctly predicts. Additionally, as future work, we would like to test the system clinically on participants with pathological gait.

Acknowledgments

This work is supported by ERC through Project 833296 (EAR), EPSRC grants EP/Y035925/1, and EP/S023046/1, and Atos.

REFERENCES

[1] T. Wu and Y. Zhao, "Associations between functional fitness and walking speed in older adults," *Geriatric Nursing*, vol. 42, 2021.

[2] A. Mirelman, P. Bonato, R. Camicioli, T. D. Ellis, N. Giladi, J. L. Hamilton, C. J. Hass, J. M. Hausdorff, E. Pelosin, and Q. J. Almeida, "Gait impairments in Parkinson's disease," *The Lancet Neurology*, 2019.

[3] S. M. Kim, D. H. Kim, Y. Yang, S. W. Ha, and J. H. Han, "Gait Patterns in Parkinson's Disease with or without Cognitive Impairment," *Dementia and Neurocognitive Disorders*, vol. 17, no. 2, pp. 57–65, Jun. 2018.

[4] C. D. Johnson, A. S. Tenforde, J. Outerleys, J. Reilly, and I. S. Davis, "Impact-Related Ground Reaction Forces Are More Strongly Associated With Some Running Injuries Than Others," *The American Journal of Sports Medicine*, vol. 48, no. 12, pp. 3072–3080, Oct. 2020.

[5] A. V. Wiik, A. Aqil, M. Brevadt, G. Jones, and J. Cobb, "Abnormal ground reaction forces lead to a general decline in gait speed in knee osteoarthritis patients," *World Journal of Orthopedics*, vol. 8, 2017.

[6] G. M. Rozanski, J. S. Wong, E. L. Inness, K. K. Patterson, and A. Mansfield, "Longitudinal change in spatiotemporal gait symmetry after discharge from inpatient stroke rehabilitation," *Disability and Rehabilitation*, vol. 42, no. 5, pp. 705–711, 2020.

[7] Y. Zhuang, J. Gong, D. C. Kerrigan, B. C. Bennett, J. Lach, and S. Russell, "Gait tracker shoe for accurate step-by-step determination of gait parameters," in *2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2016.

[8] B. Anderson, M. Shi, V. Y. F. Tan, and Y. Wang, "Mobile Gait Analysis Using Foot-Mounted UWB Sensors," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, 2019.

[9] F. A. Storm, C. J. Buckley, and C. Mazzà, "Gait event detection in laboratory and real life settings: Accuracy of ankle and waist sensor based methods," *Gait & Posture*, vol. 50, pp. 42–46, Oct. 2016.

[10] A. Baghdadi, L. A. Cavuoto, and J. L. Crassidis, "Hip and Trunk Kinematics Estimation in Gait Through Kalman Filter Using IMU Data at the Ankle," *IEEE Sensors Journal*, vol. 18, 2018.

[11] N. S. Erdem, C. Ersoy, and C. Tunca, "Gait Analysis Using Smartwatches," in *2019 IEEE 30th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC Workshops)*, Sep. 2019, pp. 1–6.

[12] F. Kluge, Y. Brand, E. Micó-Amigo, S. Bertuletti, I. D'Ascanio, E. Gazit, T. Bonci, C. Kirk, A. Küderle, L. Palmerini *et al.*, "Real-world gait detection using a wrist-worn inertial sensor: validation and comparison with the lower-back position." *JMIR Formative Research*, 2023.

[13] D. Jarchi, B. Lo, C. Wong, E. Ieong, D. Nathwani, and G.-Z. Yang, "Gait analysis from a single ear-worn sensor: Reliability and clinical evaluation for orthopaedic patients," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 8, pp. 882–892, 2015.

[14] Y. Diao, Y. Ma, D. Xu, W. Chen, and Y. Wang, "A novel gait parameter estimation method for healthy adults and postoperative patients with an ear-worn sensor," *Physiological Measurement*, 2020.

[15] A.-K. Seifer, E. Dorschky, A. Küderle, H. Moradi, R. Hannemann, and B. M. Eskofier, "EarGait: Estimation of Temporal Gait Parameters from Hearing Aid Integrated Inertial Sensors," *Sensors*, vol. 23, 2023.

[16] N. Zahradka, K. Verma, A. Behboodi, B. Bodt, H. Wright, and S. C. K. Lee, "An Evaluation of Three Kinematic Methods for Gait Event Detection Compared to the Kinetic-Based 'Gold Standard'," *Sensors*, vol. 20, 2020.

[17] C. Tudor-Locke, S. W. Ducharme, E. J. Aguiar, J. M. Schuna, T. V. Barreira, C. C. Moore, C. J. Chase, Z. R. Gould, M. A. Amalbert-Birriel, J. Mora-Gonzalez *et al.*, "Walking cadence (steps/min) and intensity in 41 to 60-year-old adults: the cadence-adults study," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 17, pp. 1–10, 2020.

[18] A. P. J. Zanardi, E. S. da Silva, R. R. Costa, E. Passos-Monteiro, I. O. Dos Santos, L. F. M. Kruel, and L. A. Peyré-Tartaruga, "Gait parameters of parkinson's disease compared with healthy controls: a systematic review and meta-analysis," *Scientific reports*, vol. 11, no. 1, p. 752, 2021.

[19] "Gait cycle wiki commons," https://commons.wikimedia.org/wiki/File:GaitCycle_by_JaquelinPerry.jpg, accessed: 2024-09-10.

[20] R. V. Briani, M. F. Pazzinatto, M. C. Waiteman, D. de Oliveira Silva, and F. M. de Azevedo, "Association between increase in vertical ground reaction force loading rate and pain level in women with patellofemoral pain after a patellofemoral joint loading protocol," *The Knee*.

[21] K. Karamanidis, A. Arampatzis, and G.-P. Brüggemann, "Symmetry and reproducibility of kinematic parameters during various running techniques," *Medicine and Science in Sports and Exercise*, 2003.

[22] J. L. McCrory, S. C. White, and R. M. Lifeso, "Vertical ground reaction forces: Objective measures of gait following hip arthroplasty," *Gait & Posture*, vol. 14, no. 2, pp. 104–109, Oct. 2001.

[23] C. Williams and C. Rasmussen, "Gaussian processes for regression," *Advances in neural information processing systems*, vol. 8, 1995.

[24] X. Jiang, C. Napier, B. Hannigan, J. J. Eng, and C. Menon, "Estimating Vertical Ground Reaction Force during Walking Using a Single Inertial Sensor," *Sensors (Basel, Switzerland)*, vol. 20, 2020.

[25] "Fermion: MPU-6050 6 DOF Sensor (Breakout)," https://www.dfrobot.com/product-880.html, 2024.

[26] N. Zahradka, K. Verma, A. Behboodi, B. Bodt, H. Wright, and S. C. Lee, "An evaluation of three kinematic methods for gait event detection compared to the kinetic-based 'gold standard'," *Sensors*, vol. 20, 2020.

[27] A. M. Jones and J. H. Doust, "A 1% treadmill grade most accurately reflects the energetic cost of outdoor running," *Journal of Sports Sciences*, vol. 14, no. 4, pp. 321–327, Aug. 1996.

[28] D. Jarchi, C. Wong, R. M. Kwasnicki, B. Heller, G. A. Tew, and Guang-Zhong Yang, "Gait Parameter Estimation From a Miniaturized Ear-Worn Sensor Using Singular Spectrum Analysis and Longest Common Subsequence," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 4, pp. 1261–1273, Apr. 2014.

[29] L. Atallah, A. Wiik, G. G. Jones, B. Lo, J. P. Cobb, A. Amis, and G.-Z. Yang, "Validation of an ear-worn sensor for gait monitoring using a force-plate instrumented treadmill," *Gait & Posture*, vol. 35, 2012.

[30] L. A. Cramer, M. A. Wimmer, P. Malloy, J. A. O'Keefe, C. B. Knowlton, and C. Ferrigno, "Validity and reliability of the insole3 instrumented shoe insole for ground reaction force measurement during walking and running," *Sensors*, vol. 22, no. 6, p. 2203, 2022.

[31] V. Mollyn, R. Arakawa, M. Goel, C. Harrison, and K. Ahuja, "Imuposer: Full-body pose estimation using imus in phones, watches, and earbuds," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23, 2023.

[32] A. Ferlini, D. Ma, R. Harle, and C. Mascolo, "EarGate: Gait-based User Identification with In-ear Microphones," in *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, Oct. 2021, pp. 337–349.

[33] C. Álvarez-Aparicio, Á. M. Guerrero-Higueras, M. Á. González-Santamarta, A. Campazas-Vega, V. Matellán, and C. Fernández-Llamas, "Biometric recognition through gait analysis," *Scientific Reports*, 2022.

[34] R. T. Shahar and M. Agmon, "Gait Analysis Using Accelerometry Data from a Single Smartphone: Agreement and Consistency between a Smartphone Application and Gold-Standard Gait Analysis System," *Sensors*, vol. 21, no. 22, p. 7497, Jan. 2021.

[35] L. Pepa, F. Verdini, and L. Spalazzi, "Gait parameter and event estimation using smartphones," *Gait & Posture*, vol. 57, pp. 217–223, Sep. 2017.

[36] F. Lin, A. Wang, Y. Zhuang, M. R. Tomita, and W. Xu, "Smart Insole: A Wearable Sensor Device for Unobtrusive Gait Monitoring in Daily Life," *IEEE Transactions on Industrial Informatics*, vol. 12, 2016.

[37] X. Liu, X. Zhang, B. Zhang, B. Zhou, Z. He, and T. Liu, "An IMU-Based Ground Reaction Force Estimation Method and Its Application in Walking Balance Assessment," *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 32, pp. 223–232, 2024.

[38] "AirPods Pro (2nd generation) - Technical Specifications," https://www.apple.com/uk/airpods-pro/specs/, 2024.

[39] M. Rahme, P. Folkeard, and S. Scollie, "Evaluating the accuracy of step tracking and fall detection in the starkey livio artificial intelligence hearing aids: a pilot study," *American journal of audiology*, vol. 30, no. 1, pp. 182–189, 2021.

[40] C. P. Burgos, L. Gärtner, M. A. G. Ballester, J. Noailly, F. Stöcker, M. Schönfelder, T. Adams, and S. Tassani, "In-Ear Accelerometer-Based Sensor for Gait Classification," *IEEE Sensors Journal*, vol. 20, no. 21, pp. 12 895–12 902, Nov. 2020.

[41] J. Prakash, Z. Yang, Y.-L. Wei, and R. R. Choudhury, "Stear: Robust step counting from earables," in *Proceedings of the 1st International Workshop on Earable Computing*, 2019, pp. 36–41.

[42] N. Jiang, T. Sim, and J. Han, "Earwalk: towards walking posture identification using earables," in *Proceedings of the 23rd Annual International Workshop on Mobile Computing Systems and Applications*, 2022, pp. 35–40.

[43] B. Lo, L. Atallah, O. Aziz, M. El ElHew, A. Darzi, and G.-Z. Yang, "Real-Time Pervasive Monitoring for Postoperative Care," in *4th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007)*, S. Leonhardt, T. Falck, and P. Mähönen, Eds. Springer, 2007.