

VIBRUN: Real-time Unobtrusive Gait Analysis for Treadmill Running via Footstep Vibrations

TIANHAO WU, University of Georgia, USA and University of Tennessee, Knoxville, USA

YI WU, University of Oklahoma, USA

BIBEK POUDEL, University of Tennessee, Knoxville, USA

SYED IRFAN ALI MEERZA, University of Tennessee, Knoxville, USA

RAJASI GORE ATHAWALE, Independent Researcher, USA

WEIZI LI, University of Tennessee, Knoxville, USA

ZAN GAO, University of Tennessee, Knoxville, USA

ÇAĞDAŞ KARATAŞ, Izmir Katip Çelebi University, Turkey

JIAN LIU, University of Georgia, USA and University of Tennessee, Knoxville, USA

Accurate and real-time gait analysis is essential for enhancing performance and reducing injury risks in treadmill running. In this paper, we introduce VIBRUN, an unobtrusive gait analysis system that estimates key physiological metrics, such as cadence, ground contact time, stride time, center of pressure, and plantar pressure distribution, through footstep vibrations captured by low-cost treadmill-mounted sensors. Leveraging advanced multi-task transformer models, our system offers a robust, real-time solution to monitor and analyze running biomechanics without requiring intrusive wearable devices. This approach enables seamless integration into virtual sports, gaming platforms, and immersive exercise environments, enhancing the running experience by providing personalized feedback. By offering precise biomechanical insights in real-time, VIBRUN paves the way for future applications in virtual sports, gamified fitness, and interactive training programs, empowering users to engage more effectively in their training sessions while improving overall performance and reducing injury risks. Extensive evaluations with 17 participants across varied treadmill running scenarios demonstrate VIBRUN's accuracy in real-time gait analysis. For instance, VIBRUN achieves a mean error of 28.8 ms in ground contact time and a distance of 13.66 mm in the center of pressure, among other measured metrics, highlighting its precise performance across multiple gait parameters.

1 INTRODUCTION

Running is undoubtedly one of the most popular forms of exercise worldwide. A recent report indicates that the number of runners in the United States reached 48.98 million in 2024 [1]. While outdoor running remains the most common form, indoor running is gaining popularity due to its convenience and adaptability to various environmental conditions and physical constraints, with the global treadmill market estimated to reach \$4 billion in 2024 [2]. Consequently, various online running platforms and games have been developed to make indoor running more immersive, engaging, and competitive [3–6]. Runners can even participate in online running races through virtual running platforms (e.g., Zwift [7]). To increase runner engagement, provide personalized feedback [8, 9] for performance improvement, enhance the realism of online running sessions, and prevent potential injuries, running gait analysis has become essential for monitoring key metrics such as cadence, ground contact time, and plantar pressure distribution. For example, Van Rheden *et al.* investigate multimodal feedback systems for sports wearables and highlight how different feedback modalities, such as visual, auditory, and haptic, can be used to enhance performance and technique without interfering with athletic activities [8].

Prior Work on Running Gait Monitoring. Existing running gait monitoring systems typically use wearable sensors [10, 11], such as Inertial Measurement Unit (IMU) sensors, to estimate various metrics, including step count [12], foot strike [12, 13], heading direction, speed [14], and stride length [15]. However, these systems have limited tracking capabilities and can only reconstruct certain metrics. Although smart insoles can monitor more fine-grained metrics, such as Center of Pressure (COP) and Plantar Pressure Distribution (PPD), they are often criticized for their high cost; for example, the Tekscan F-Scan64 pressure mapping system costs over \$8,000 [16],

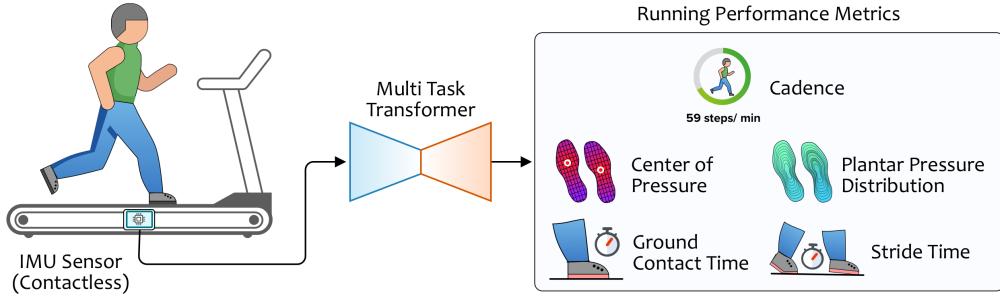


Fig. 1. VibRun: An unobtrusive gait monitoring system using IMU sensor and a Multi-Task Transformer model for deriving running-related metrics such as cadence, ground contact time, stride time, center of pressure, and plantar pressure distribution.

and the Noraxon smart insole system exceeds \$10,000 [17]. Additionally, these wearable-based approaches require attaching sensors to the runner’s body, which inevitably affects mobility and comfort. Furthermore, several unobtrusive sensor-based gait monitoring systems have been developed, including acoustic-based [18–21], vision-based [22–24], and radar/lidar-based [25–27]. However, these systems often employ expensive hardware and software [28, 29], which can be inconvenient to set up. To the best of our knowledge, no previous non-intrusive running gait monitoring system offers comprehensive feedback across a wide range of running metrics in an unobtrusive, low-cost, and easily deployable manner.

System Objectives and Challenges. To overcome the aforementioned limitations, we develop an unobtrusive gait analysis system for treadmill running that can monitor five important running metrics, including Cadence, Ground Contact Time (GCT), Stride Time (ST), Center of Pressure (COP), and Plantar Pressure Distribution (PPD). These metrics can provide personalized feedback to runners, aiding in enhancing training performance, preventing potential injuries, and increasing the realism of virtual platforms. While commercial wearables can estimate basic running metrics such as cadence or pace, they fall short in delivering detailed biomechanical insights like plantar pressure distribution and center of pressure, which are critical for assessing running technique, identifying asymmetries, and preventing injuries. VibRun bridges this gap by offering real-time feedback on these advanced metrics without requiring specialized footwear or lab-based equipment. This accessibility is especially valuable for fitness enthusiasts aiming to optimize performance, track progress, or reduce injury risk, making high-quality gait analysis feasible in regular treadmill workouts. We observe that while a runner is running on a treadmill, the foot striking the treadmill surface generates vibrations that propagate through the treadmill and can be captured by an IMU sensor. We chose vibration sensing with an IMU sensor because vibration signals are resilient to ambient noise (e.g., background sounds), non-intrusive to the runner (requiring no wearable sensors), and cost-effective. As the runner varies their speed and strike patterns, the resulting vibration signals will exhibit corresponding patterns that are useful during running gait analysis. However, implementing such a system faces several challenges:

(1) *Mapping Vibration Signals to Running Metrics:* Vibration signals provide only non-directive and coarse-grained motion sensing data. The mapping between this type of data and running metrics remains unexplored.

(2) *Per-foot & Real-time Analysis:* The system must accurately differentiate the vibrations caused by each foot and be both cost-effective and computationally efficient to enable real-time gait analysis.

(3) *Multi-task Design:* The system must be capable of multitasking and predicting necessary metrics. Given that the metrics are on different scales, it is crucial to design a loss function that accounts for their variations.

Unobtrusive Gait Analysis via Footstep Vibrations. To tackle these challenges, we propose VibRun, an unobtrusive running gait analysis system. As illustrated in Fig. 1, an IMU sensor is attached to the side of the treadmill to capture the vibration signals generated by foot strikes. These signals are then processed by a multi-task transformer model to reconstruct five key running performance metrics. The design of VibRun is specifically

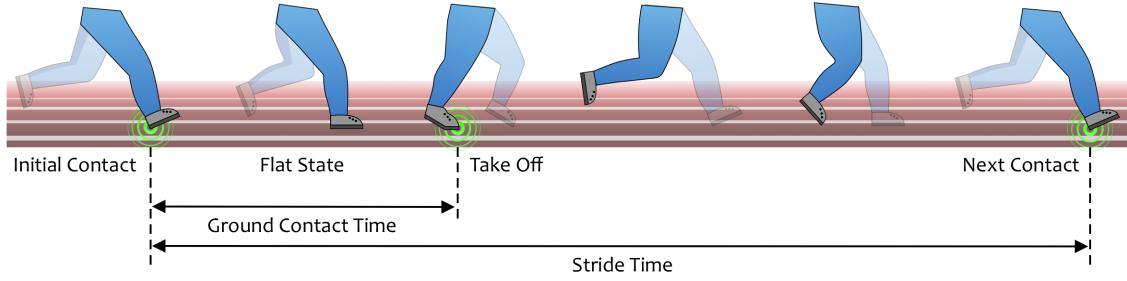


Fig. 2. Illustration of Ground Contact Time and Stride Time.

optimized to maintain a minimal level of intrusiveness for runners while remaining robust against environmental noise associated with running activities. More specifically, we adopt the traditional encoder-decoder architecture for the multi-task transformer model, incorporating multi-head attention and encoder-decoder attention to capture the temporal dependencies within the vibration signals. To reduce the computational cost, we use a single encoder to generate a shared feature representation for all tasks, while two separate decoders are employed to reconstruct COP and PPD, respectively. Given the reconstructed COP, we develop a signal process algorithm to further derive the Cadence, GCT, and ST of the runner. To further enhance the system's adaptability across different users and running styles, we integrate an unsupervised domain adaptation (UDA) module into our framework. This module aligns the feature representations between seen and unseen users without requiring any labeled data (e.g., ground truth from the insole sensor) for new-user calibration, allowing ViBRUN to generalize to new individuals and maintain reliable performance in real-world treadmill running scenarios. The overall system is cost-effective and computationally efficient, enabling seamless deployment on a Raspberry Pi to perform real-time analysis. Our contributions are summarized as follows:

- We propose ViBRUN, the first unobtrusive, running-gait-analysis system for treadmill running by leveraging footstep vibrations. ViBRUN can reconstruct five running metrics: Cadence, Ground Contact Time (GCT), Stride Time (ST), Center of Pressure (COP), and Plantar Pressure Distribution (PPD). These metrics can serve as personalized feedback to runners, helping to improve running efficiency, boost engagement in virtual sports, and prevent potential injuries.
- ViBRUN maintains a minimal level of obtrusiveness for runners while providing comprehensive feedback across five key running metrics. ViBRUN is also cost-effective and robust against environmental noise.
- We develop a multi-task transformer model that maps coarse-grained vibration signals to five running metrics. ViBRUN can be integrated into any treadmills and can perform real-time gait analysis.
- Extensive evaluations with 17 participants demonstrate ViBRUN's accuracy in real-time running gait analysis, with mean errors of 1.69 steps per minute for Cadence, 29 ms for GCT, 44 ms for ST, an L2 distance of 13.66 mm for COP, and a mean cosine similarity of 0.879 for PPD.

2 SYSTEM OBJECTIVES AND CHALLENGES

2.1 System Objectives

In this work, we aim to derive five key metrics using ViBRUN to enhance performance, safety, and engagement in virtual sports, gamified fitness, and interactive training. These metrics provide real-time feedback to optimize efficiency, synchronize virtual avatars, and reduce injury risks, enabling more effective and immersive training sessions.

- **Cadence.** Cadence, the number of steps per minute, is a key factor in running performance that directly influences the speed and the realism of virtual sports environments. In gamified fitness and interactive training

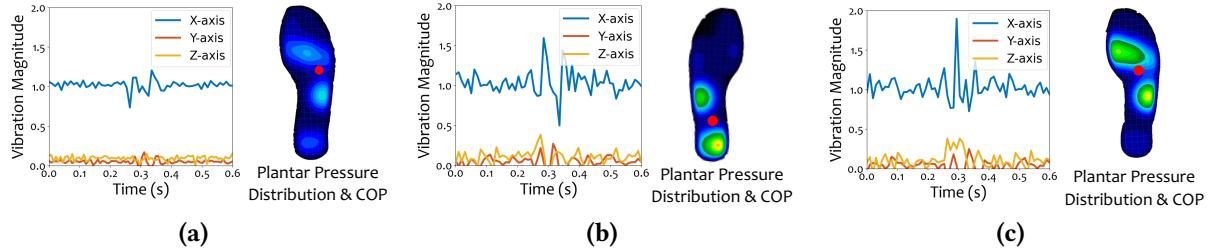


Fig. 3. Examples of footstep vibrations illustrating the relationship between PPD and COP with the vibrations detected by the IMU sensor.

programs, maintaining an optimal cadence not only improves efficiency but also enhances the immersive experience [30, 31].

- **Ground Contact Time (GCT).** GCT refers to the duration during which a runner's foot is in contact with the ground, from initial contact to toe-off, as shown in Fig. 2. Monitoring GCT is crucial because prolonged ground contact can lead to inefficiencies in running form, increasing the metabolic cost of running and slowing down the runner. Additionally, excessive GCT can result in greater impact forces on the joints, particularly the knees and ankles, elevating the risk of injury [32].
- **Stride Time (ST).** ST is the total duration required to complete a single stride, including both the time spent in the air and the time spent on the ground [33], as shown in Fig. 2. Longer strides might seem beneficial for speed, but they can increase impact forces on the knees, raising the risk of injuries such as patellofemoral pain syndrome. Conversely, overly short strides might limit speed potential. Monitoring and optimizing ST helps ensure that runners maintain an efficient stride length and frequency, ultimately reducing injury risk and improving performance [34, 35].
- **Center of Pressure (COP).** COP represents the centroid of the pressures over the surface of the area in contact with the ground. It's closely related to gait performance and foot & ankle function. A more lateral position of the COP results in a smaller frontal plane moment arm of the force around the ankle, thereby decreasing the rearfoot tendency toward eversion [36]. Additionally, the trajectory of COP could be utilized to diagnose foot pronation and supination [37].
- **Plantar Pressure Distribution (PPD).** PPD measures the force distribution across the sole during foot contact with the ground, offering valuable insights into biomechanics. By analyzing PPD, runners can identify imbalances that may lead to discomfort or injury, particularly in the foot, ankle, and knee [38]. In this work, we aim to derive pressure levels in eight foot regions—Arch, Hallux, Heel (two regions), Metatarsal bones (three regions), and Toes—to optimize running form, enhance performance, and minimize injury risks during training.

2.2 Challenges

Realizing a system to track the aforementioned running metrics raises several practical challenges:

- (1) *Running Gait Analysis Using Unobtrusive IMU Sensors:* Unlike existing studies that rely on high-fidelity sensors, such as microphones and cameras, or require attaching sensors to the runner's body, our system employs only IMU sensors mounted on the treadmill. To validate whether these sensors can capture running activities, we conducted preliminary experiments using a 9-axis IMU sensor connected to a Raspberry Pi 4, mounted on the side frame of multiple treadmill models sourced from different vendors. We asked a participant to run on the treadmill with different foot landing patterns: forefoot landing, where the COP is at the front part of the foot as illustrated in Fig. 3 (a) and (c), and rear-foot landing, where the COP is at the latter part of the foot, as illustrated in Fig. 3 (b).

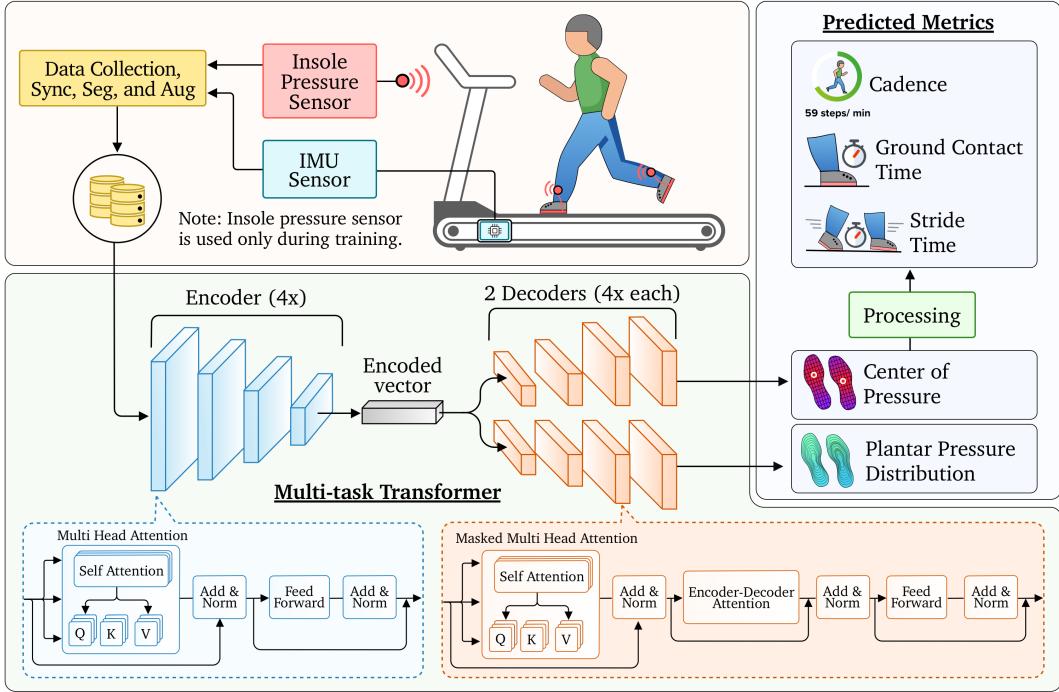


Fig. 4. ViBRUN system architecture. When a user runs on a treadmill, ViBRUN captures data in an unobtrusive manner through an IMU attached to the Treadmill. The insole pressure sensors are only used to collect ground truth for training. The segmentation and synchronization of captured data is performed on a Raspberry Pi, after which the processed data is augmented (only during training) and fed into a multi-task transformer model, that includes a shared encoder and two task-specific decoders to predict the center of pressure and plantar pressure distribution, respectively. For a comprehensive view of running performance, further a binary conversion algorithm is applied to the center of pressure prediction to obtain metrics such as cadence, ground contact time, and stride time.

While we observed distinct vibrations during the run, the IMU sensors provide only non-directive, coarse-grained motion data. The challenge of mapping this IMU data to detailed running metrics remains unexplored.

(2) *Per-foot & Real-time Analysis*: Our system must differentiate between the vibrations generated by each foot and reconstruct the five running metrics accordingly. However, as illustrated in Fig. 3 (b) and (c), the vibrations generated by the left and right feet exhibit a high level of similarity, posing a significant challenge in identifying which foot is in contact with the ground. Furthermore, the system must be both cost-effective and computationally efficient to enable real-time gait analysis.

(3) *Multi-task Design*: We need to design a multi-task model that can regress all required metrics simultaneously. We must ensure that the learned representations are sufficiently generalized to be applicable across different metrics, while also developing specialized downstream decoders for each metric. Additionally, since the metrics vary in scale, it is critical to design an appropriate loss function that accounts for these differences.

3 SYSTEM OVERVIEW

The overall system architecture is illustrated in Fig. 4. During the training phase, we collect two types of data: pressure sensing data from a pair of commercial insole pressure sensors (Noraxon Ultium Insoles [17]) to obtain the ground truth for COP and PPD, and vibration data from an IMU sensor (i.e., MPU 6050). The IMU sensor

is integrated with a Raspberry Pi and mounted on the side of the treadmill. It is important to note that the pressure sensors are only used for ground truth acquisition during the training phase. In the testing phase, VIBRUN can reconstruct the five running metrics using only the vibration data. We first perform *Data Processing, Synchronization, Segmentation, and Augmentation*. We synchronize the pressure sensing data and the vibration data, then segment both data into separate data frames and do augmentation. The augmented vibration data is used as input for the *Multi-task Transformer* model, which is designed to reconstruct COP and PPD based solely on vibration signals.

We employ a multi-task transformer model for several key reasons. First, the shared encoder in the transformer allows us to capture high-level feature representations common to both tasks, maximizing the efficiency of feature extraction. This shared representation reduces the computational complexity of the system by avoiding redundant feature extraction for each task. Furthermore, the inherent design of transformers, with their attention mechanisms, is well-suited for capturing temporal dependencies within the vibration data, which is critical for accurate gait analysis. The use of a multi-task setup ensures that both COP and PPD are reconstructed simultaneously, allowing the model to leverage the underlying correlation between these two metrics for improved performance. Each task has its own dedicated decoder, which ensures that while they share a common representation, task-specific nuances are captured during reconstruction.

Specifically, the transformer architecture incorporates multiple *Multi-Head Attention* modules to focus on different segments of the input sequence and capture complex temporal patterns within the vibration signals. Each attention module consists of *Self-Attention*, *Add & Norm*, and *Feed Forward* layers to ensure robust learning. In the decoder stage, an additional *Encoder-Decoder Attention* module helps emphasize relevant parts of the encoded input sequence, further refining the reconstruction of COP and PPD. Detailed explanations of these components will be provided in Section 4.3. Finally, through *Post-Signal Analysis*, the reconstructed COP is converted to a binary signal indicating whether the foot is on the ground. This binary information is then used to derive the remaining three metrics: Cadence, GCT, and ST. During the testing phase, the vibration data will be segmented and input into the pre-trained multi-task transformer model, which is deployed on the Raspberry Pi. This setup allows for real-time inference of the five running metrics, enabling immediate analysis and feedback based on the incoming data.

4 DESIGN OF VIBRUN

4.1 Physical Vibration Propagation

During running on a treadmill, interactions between a runner’s feet and the moving belt generate vibrations. As illustrated in Fig. 5(a), each foot strike creates vibration waves that originate at its COP. These waves propagate through the treadmill belt and frame, reflecting from boundaries and interfering with other waves before being detected by the onboard IMU sensor. To characterize how running metrics influence the captured vibration signals, we consider two primary factors:

Temporal Factors: Metrics such as GCT, ST, and Cadence directly influence the temporal patterns observed in the vibration signals. Different phases of the foot’s interaction with the treadmill belt, such as contact and take-off (shown in Fig. 2), produce distinct signatures in vibration amplitude. Specifically, during foot contact, short vibration impulses with high amplitude are observed, whereas during the period between previous take-off and next foot contact (when the runner is briefly airborne), the overall amplitude is reduced. Further, ST creates longer repeating cycles in the signal corresponding to consecutive impacts of the same foot, whereas Cadence determines the overall frequency of vibration impulses, regardless of which foot generates them.

Force and Spatial Factors: To illustrate the relationship between impact-force magnitude and its location on the resulting vibrations, we approximate the treadmill belt as the spring–mass–damper system [39] shown in Fig. 5(b) where the springs represent elastic stiffness (energy storage) and dampers represent energy dissipation

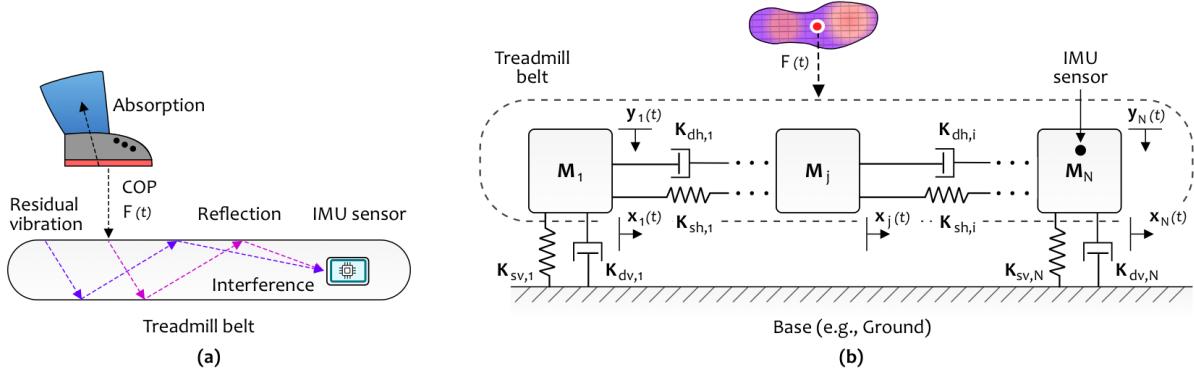


Fig. 5. Vibration propagation on a treadmill belt. (a) Physical process: A foot strike applies a vertical force at the COP. The resulting vibration is partially absorbed by the shoe; the remainder travels along the belt, reflects at the boundaries, interacts with residual vibration waves, and is ultimately captured by the IMU sensor; (b) System model: The belt is segmented into N discrete masses (M_1, \dots, M_N) along its length and represented as a spring–mass–damper system. Adjacent masses are coupled horizontally by springs $K_{sh,i}$ and dampers $K_{dh,i}$ ($i = 1, \dots, N - 1$). The first and last masses are connected to a rigid base (representing supports that anchor the treadmill to the ground) through vertical spring–damper pairs $K_{sv,1}, K_{dv,1}$ and $K_{sv,N}, K_{dv,N}$. When a foot strike excites the mass M_j with force $F(t)$, it produces horizontal displacements $x_j(t)$ for all j ($j = 1, \dots, N$) masses and vertical displacements $y_1(t), y_N(t)$ at the end masses. An IMU sensor is mounted on the right-hand end mass M_N .

(material damping). The belt is segmented into N discrete masses ($M_1 \dots M_N$) along its length, with adjacent masses coupled horizontally by springs $K_{sh,i}$ and dampers $K_{dh,i}$, allowing horizontal displacements $x_i(t)$. The first and last masses, where the treadmill is supported by ground, are attached to the base by vertical spring–damper pairs $K_{sv,1}, K_{dv,1}$ and $K_{sv,N}, K_{dv,N}$, permitting vertical displacements $y_1(t)$ and $y_N(t)$, respectively. A foot strike is modeled as a time-varying force $F(t)$ applied at the COP, exciting the mass segment M_j directly below it.

Within this system model, we focus on the horizontal dynamics of the treadmill belt for simplicity, i.e., the horizontal displacement measured at the sensor mounted on the final mass M_N is jointly influenced by (i) the magnitude of impact force $F(t)$, (ii) the strike location j (i.e., which mass segment M_j is excited), and (iii) the spring–mass–damper parameters $\{K_{sh,i}, K_{dh,i}, \dots\}$ that characterize the belt. Let $x_j(t)$ denote the horizontal displacement of segment j , the resulting equations of motion are

- **Boundary segments ($j = 1, N$):**

$$M_1 \ddot{x}_1 = K_{sh,1}(x_2 - x_1) + K_{dh,1}(\dot{x}_2 - \dot{x}_1) + f_1(t), \quad (1)$$

$$M_N \ddot{x}_N = K_{sh,N-1}(x_{N-1} - x_N) + K_{dh,N-1}(\dot{x}_{N-1} - \dot{x}_N) + f_N(t), \quad (2)$$

- **Interior segments ($2 \leq j \leq N - 1$):**

$$M_j \ddot{x}_j = K_{sh,j-1}(x_{j-1} - x_j) + K_{sh,j}(x_{j+1} - x_j) + K_{dh,j-1}(\dot{x}_{j-1} - \dot{x}_j) + K_{dh,j}(\dot{x}_{j+1} - \dot{x}_j) + f_j(t). \quad (3)$$

Here $\dot{x}_j(t)$ denotes velocity, $\ddot{x}_j(t)$ denotes acceleration, and $f_j(t)$ denotes the external horizontal force acting on the mass segment M_j . For a single foot-strike landing on segment j we set $f_j(t) = F(t)$ and assign zero force to all other segments. Because the dynamics in Eqs. (1)–(3) are coupled, a force applied at M_j propagates through the chain of spring–damper pairs. Thus, vibrations initiated far from the sensor location (M_N) still reach M_N after a distance-dependent delay and with an amplitude that diminishes according to the cumulative stiffness [39] and damping encountered along the path.

Note that this simplified representation provides an intuitive illustration of how changes in force location alter the timing and amplitude of the vibration signals reaching the sensor. In a real system, the dynamic response (due

to the interplay of masses, springs, and dampers) also modifies the shape and frequency of propagating signals. Further, our system model divides the treadmill into N segments along its length, but a treadmill belt is a 2D surface, and vibrations propagate along both length and breadth, creating more complex interactions. Additional factors such as varying treadmill properties (weight and belt thickness), nonlinearities (absorption, reflection, and interference of multiple simultaneous vibrations), and differing runner biomechanics (such as weight, height, and running form) make it difficult to accurately derive the running metrics using model-based approaches. Hence in practice, obtaining running metrics from the analysis of vibration signals using mathematical models is highly challenging. Therefore, we leverage machine learning to learn the mapping directly from the measured vibration signals to the target running metrics.

4.2 Data Collection, Synchronization, Segmentation and Augmentation

We collect data from two main sources: pressure sensing data from a pair of Noraxon Ultium insoles and vibration signals (i.e., accelerometer data) from the developed VIBRUN, which consists of an MPU 6050 IMU sensor connected to a Raspberry Pi. The pressure sensing data is used solely as ground truth for model training and testing.

Synchronizing the data streams is crucial for accurate analysis. To achieve this, participants are instructed to stand on their right foot and then step forcefully on the treadmill using the left foot at the start of data collection. This action generates a distinctive spike in both sensor readings, facilitating precise alignment and synchronization of the data streams. To prepare the data for further processing, we use the Standard Score (Z-score) normalization method [40], which standardizes the data by expressing each data point in terms of standard deviations from the mean. This normalization step adjusts the data to have a mean of zero and a standard deviation of one, improving the consistency of the input data and enhancing model performance [41].

Then the normalized 3-axis accelerometer data is used for segmentation, while the pressure sensing data is processed using the MyoMUSCLE software [42] to obtain the groundtruth for COP and PPD. We set the sampling rate for the COP and PPD data at 250 Hz, which preserves sufficient detail of the running gait. The sampling rate for the accelerometer data is fixed at 100 Hz, ensuring consistent and accurate capture of vibration signals for analysis. All data streams are segmented into 3-second windows. To enhance data diversity and improve model robustness, we apply signal-based data augmentation techniques to the training set. Specifically, we inject Gaussian noise into the accelerometer data and apply random amplitude scaling within a $\pm 10\%$ range. This process effectively triples our training dataset size, with each original sample having two augmented counterparts. All samples in the original training set undergo both augmentation techniques, ensuring comprehensive coverage of potential signal variations. The Multi-Task Transformer model takes a 3-second segment of accelerometer data as input and outputs the continuous COP and PPD values over the corresponding 3-second period.

4.3 Multi-task Transformer Model Design

The overview of our multitask model design is shown in Fig. 4. We adapt the encoder-decoder transformer architecture [43], which has demonstrated its ability to capture long-range temporal dependencies and complex patterns in time series data [44]. Motivated by the multi-task transformer’s success in computer vision [45], we modify the transformer to include a single encoder that is shared across two task-specific decoders. Both COP and PPD require the capture of spatial and temporal features: the COP reflects the dynamic location of the pressure center on the foot over time; meanwhile, the PPD captures how force is distributed and spread across the foot at each moment, involving both spatial patterns and their changes over time. Consequently, a shared encoder is sufficient to extract both feature types. The Transformer’s positional encoding enables effective temporal modeling, while the multi-head attention mechanism ensures robust feature extraction across both spatial and temporal domains. This design simplifies the architecture, reduces computational costs, and maintains high performance.

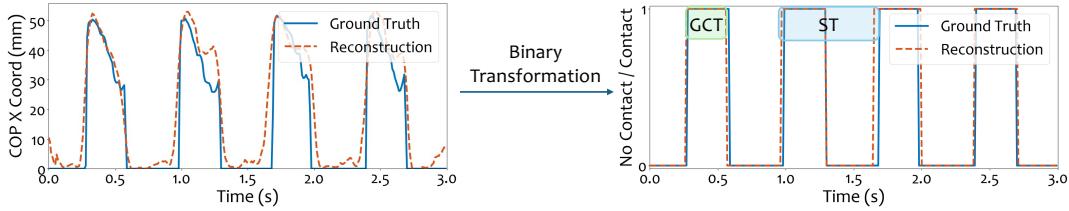


Fig. 6. Illustration of post-signal analysis.

Specifically, our model consists of four stacked encoder layers and two decoders, each decoder is also composed of four stacked decoder layers. Both encoder and decoder layers utilize the multi-head attention mechanism proposed in the original transformer which, unlike windowed or chunked attention commonly found in transformers for time series, allows us to capture the global dependencies across the entire input. In this multi-head attention mechanism, each attention head is represented by Query (Q), Key (K), and Value (V) matrices. These matrices are used to compute attention scores and aggregate information from different positions in the sequence. After concatenating the outputs from all attention heads, the resulting contextual embeddings are produced. Following this, residual connections and layer normalization are applied to the embeddings before they are fed into a two-layer feed-forward neural network. Additionally, in the decoder layers, the multi-head attention mechanism applies a “subsequent” mask to ensure that predictions at a given time step depend only on preceding time steps. The output from this self-attention mechanism is subsequently processed by an encoder-decoder attention mechanism, allowing the decoder to focus on relevant segments of the encoded input. Finally, two linear layers are used after each decoder: one to produce COP predictions and the other to generate PPD predictions. Details of the model training settings and hyperparameters can be found in Section 5.2.

4.4 Initial Calibration via Unsupervised Domain Adaptation

To enable the usability across different individuals, our system adopts an initial calibration phase that does not require manual labeling of ground truth data for new users. Instead, we utilize UDA [46] to align the model’s performance between a labeled source domain (data from previously seen subjects) and an unlabeled target domain (the new subject). This calibration phase requires only a few minutes of IMU data collected across different running speeds. Our UDA approach follows the Domain-Adversarial Neural Network (DANN) framework, where a shared Transformer encoder is jointly optimized for both the primary reconstruction task (COP and PPD prediction) and a domain discrimination task. The encoder is trained to extract domain-invariant features by minimizing reconstruction loss on source data while simultaneously maximizing confusion for a domain classifier that attempts to distinguish between source and target samples. This is achieved using a Gradient Reversal Layer (GRL) that inverts the gradient signal from the domain classifier, encouraging the encoder to produce features that are indistinguishable across domains. Our domain classifier is implemented as a lightweight two-layer MLP that maps pooled Transformer features into domain logits. Specifically, features are first projected into a 128-dimensional intermediate space, followed by a ReLU activation and a second linear projection into a 64-dimensional hidden space, and then classified into source or target domains. Once trained, the model can reconstruct target-domain COP and pressure signals with high fidelity, despite not having access to their corresponding ground truth during calibration. This enables real-time, personalized inference for new users with no manual annotation effort, fulfilling the goal of our system.

4.5 Metric Derivation through Post-Signal Analysis

During treadmill running, COP data provides critical insights into foot-ground contact dynamics. Specifically, COP values only present when the foot is in contact with the treadmill and are 0 while the foot is in the air,

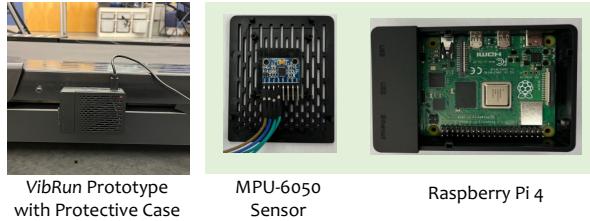


Fig. 7. Hardware implementation of VibRun showing components within the case and connections between the components.

allowing us to develop a threshold-based mechanism to generate a binary waveform indicating foot-ground contact for GCT, cadence and ST. As shown in Fig. 6, we leverage the X coordinate value of the left foot to derive the binary waveform. Note the post-signal analysis algorithm is the same for both feet. For the groundtruth data, any COP value greater than 0 is converted to 1 (indicating foot contact with the ground), while values equal to 0 remain unchanged (indicating the foot is in the air). For the reconstructed COP, we evaluate a range of potential thresholds to determine the one that provides the most accurate prediction. Specifically, we utilize the Receiver Operating Characteristic (ROC) curve on the validation dataset (details about the dataset split are provided in Section 6.1). The ROC curve is used to assess the trade-off between the true positive rate (correctly detected foot contact) and the false positive rate (incorrectly detected foot contact) across different thresholds. We select the threshold that maximizes the Area Under the Curve (AUC). To improve accuracy, we apply an outlier removal algorithm based on the duration of consecutive “1”s in the binary waveform. Since GCT cannot be as short as 20 milliseconds (with fast runners typically having values in the range of 175-200 milliseconds [47]), any spikes with such brief durations are identified as outliers and removed.

Fig. 6 illustrates an example of the binary waveform. GCT is derived as the duration between a 0-to-1 transition and an adjacent 1-to-0 transition, while ST represents the duration of a complete cycle. The cadence range for jogging and running typically falls between 100 and 200 steps per minute (SPM) [48, 49]. To calculate cadence, we apply a Fast Fourier Transform (FFT) on the reconstructed COP waveform within this range, corresponding to a per-foot frequency of approximately 0.83 Hz to 1.67 Hz. We focus on the cadence for a single foot because the data treats the left and right feet separately. By focusing on this specific frequency range, we ensure accurate cadence extraction from the binary waveform.

5 IMPLEMENTATION

5.1 Hardware Implementation

As illustrated in Fig. 7, a single IMU sensor is connected to a Raspberry Pi 4 in the prototype of VibRun. The IMU sensor, an MPU-6050 [50] featuring a 3-axis accelerometer, is used to capture running footstep vibrations. The IMU sensor is configured with a sampling rate of 100 Hz. These hardware devices communicate using the I2C (Inter-Integrated Circuit) protocol, which requires the use of SDA (Serial Data) for transmitting data and SCL (Serial Clock) for synchronizing signals. This case is mounted on one side of the treadmill, near the frame adjacent to the running belt. This placement ensures proximity to the footstrike area where vibrations are most prominent. The system can be attached using industrial-grade Velcro strips or clamps to keep it stable during running sessions, preventing unwanted movement or shifts.

5.2 Machine Learning Model Implementation

The input to the transformer model consists of 3-axis IMU data with dimensions of (300, 3). A linear layer projects this raw data to dimensions of (300, 64), which is then passed as input to the encoder. To preserve temporal information within the sequence, positional encoding is applied using sinusoidal functions [43]. Both the encoder and decoder layers employ 16 attention heads, with each head receiving an equal share of the 64-dimensional

Table 1. Training parameters and hyper-parameters for the multi-task transformer model.

Hyper-parameter	Value
Number of attention heads (h)	16
Embedding size (d_{model})	64
Feedforward dimension (d_{ff})	2048
Optimizer	Adam
Dropout rate	0.2
Learning rate initial (α)	2×10^{-4}
Batch size	128
Sequence length	300
Input dimension (IMU)	3
Output dimension (COP)	4
Output dimension (Pressure)	16
Number of epochs	1000

input, i.e., $64/16 = 4$ dimensions per head. The Query, Key, and Value (QKV) matrices have dimensions of $(300, 4)$ for each attention head. The feed-forward network, used consistently across all layers, has two stages: the first stage projects the data from 64 to 2048 dimensions, while the second stage reduces it back from 2048 to 64 dimensions. Each decoder is followed by a separate linear layer that serves as a projection head to map the output to the corresponding prediction dimension—one for COP and the other for PPD. The final linear layer for estimating COP has a dimension of 4 (the x and y coordinates for both feet). Another linear layer, used for PPD, outputs 16 values (corresponding to 8 regions on each foot).

Optimization. The model is trained for 100 epochs using a loss function that combines two components: the Mean Squared Error (MSE) between the ground truth and predicted PPD for the PPD decoder L_{PPD} , and the MSE between the ground truth and predicted COP for the COP decoder L_{COP} . The final loss is the weighted sum of the two components $L = L_{PPD} + 4 * L_{COP}$, which is then backpropagated through the network. The weight factor of 4 is used to balance differences in task complexity or error scales between the two predictions. Additionally, we employ the Adam optimizer [51] with a dropout rate of 0.2 to mitigate overfitting. The batch size is set to 128 and the learning rate is initialized at 2×10^{-4} which reduces to 2×10^{-5} at 50 epochs and 2×10^{-6} at 75 epochs. A comprehensive list of the hyperparameters used is provided in Table 1.

5.3 Real-Time Deployment

The real-time deployment of ViBRUN on a Raspberry Pi 4 is illustrated in Fig. 8. The data pipeline begins with the IMU sensor capturing acceleration signals, which are continuously collected at a sampling rate of 100 Hz. Each data packet is timestamped and segmented into fixed-length windows (e.g., 3-second windows) for real-time analysis. These windows are then fed into the pre-trained ViBRUN multi-task transformer model, which has been optimized for low-latency inference on edge devices. The model performs feature extraction and predicts two key metrics: COP and PPD. In parallel, additional running metrics such as Cadence, GCT, and ST are derived from the COP predictions.

To enable efficient deployment across multiple edge devices in different locations, a Docker container has been created, bundling the necessary dependencies, including *ExecuTorch* for model inference, custom Python scripts for signal processing, and hardware drivers for the IMU sensor. This container ensures consistent execution regardless of the specific hardware environment. The size of the pre-trained model, which uses quantization techniques to reduce memory and computational requirements without sacrificing accuracy, is 18.1 MB, allowing it to run efficiently on the limited resources. Our system can easily be extended to transmit data to nearby mobile devices, virtual reality (VR) systems, or smart treadmills for further processing. This extension would enable more advanced fitness data analysis and support interactive/adaptive training, as well as virtual esports applications.

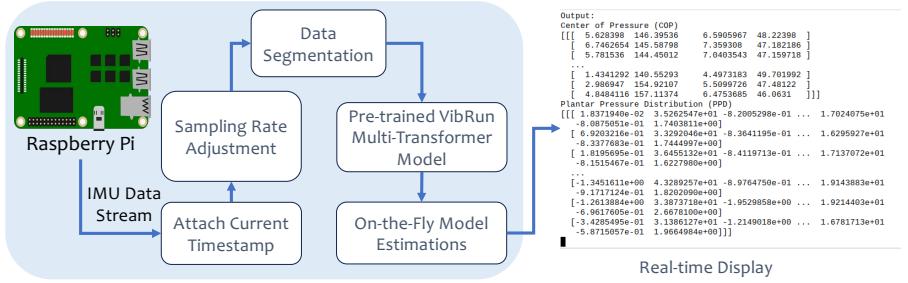


Fig. 8. Data processing steps for real-time deployment on Raspberry Pi.

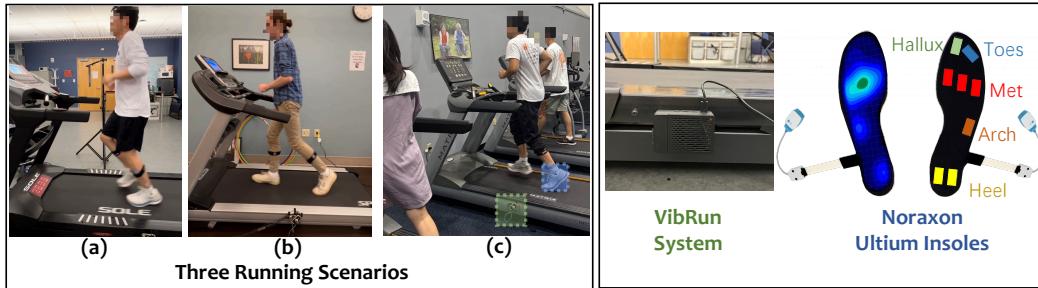


Fig. 9. Illustration of data collection across various scenarios using the VibRun system and Noraxon Ultium insoles.

6 SYSTEM EVALUATION

In this section, we outline the experimental setup for data collection, the evaluation methods used for our VibRun system, and the visualization of the reconstructed running metrics.

6.1 Experimental Setup & Data Collection

We recruited a total of 17 participants to evaluate the performance of VibRun¹. The study participants include 12 males and 5 females, with heights ranging from 5'3" to 6'1" and body weights from 103 to 210 lbs. Their physical activity levels vary, including both non-daily exercisers and weekly runners. During data collection, each participant was asked to run on a treadmill at four different speeds, ranging from 3 mph to 6 mph, with each speed session lasting approximately 6 minutes. Participants were allowed to choose their own running shoes and run in their preferred, most comfortable style. To ensure functionality and stability, VibRun, composed of a Raspberry Pi 4 and an MPU 6050 accelerometer sensor, is securely attached to the middle side of the treadmill using Velcro tape, as shown in Fig. 1. For reliable ground truth on running metrics such as PPD and COP, we use Noraxon Ultium Insoles [17], which provide 8-zone PPD data for each foot at a sampling rate of 250 Hz. To be more specific, these 8 zones are the Arch, Hallux, Heel (2 regions), Metatarsal (Met) bones (3 regions), and Toes.

To demonstrate the generalizability of VibRun, we evaluated it on various popular treadmill models, including the Sole F63, Spirit XT685, and Matrix T5X. The Matrix T5X is a commercial-grade unit equipped with the Ultimate Deck™ system, which includes a 1-inch hard-wax reversible deck and a welded steel frame, providing high structural rigidity [52]. The Spirit XT685 offers a balance between home and light-commercial use, featuring a waxed 1-inch deck and heavy steel frame for moderate rigidity [53, 54]. The Sole F63, a foldable home treadmill, includes the Cushion Flex Whisper Deck, which reduces impact forces by up to 40% compared to asphalt, resulting

¹Our study has been approved by our Institutional Review Board (IRB).

in a lower-rigidity platform [55]. All three treadmills are equipped with standard 2-ply belts, commonly found in mid- to high-end models. The Spirit XT685 uses a 2.5 mm thick belt [54], while the Matrix T5X and Sole F63 use belts approximately 2.2–2.3 mm thick [52, 56]. While each offers a similar running surface (approximately 60 × 20–22), they differ in weight and frame stability—ranging from 394 lbs (Matrix T5X) [52] to 269 lbs (Spirit XT685) [54] and 267 lbs (Sole F63) [57]. These differences ensure that our evaluation encompasses a representative range of treadmill types and real-world scenarios.

As shown in Fig. 1, our experiments were conducted in three different lab environments. The primary experiments took place in environments (a) and (b), where only a single treadmill was present in the room. Additional data was collected in a kinesiology laboratory (c), which contained multiple treadmills, for an ablation study aimed at assessing the impact of environmental vibrational noise from nearby operating treadmills (Section 6.4). It is important to note that in all experimental environments, there were frequently people walking in the room, inevitably creating footstep vibrations that could potentially influence the system.

Additionally, we also conducted a longitudinal study in which data was collected from the same subject with a time gap of approximately one month. This enabled us to assess the system’s stability over time. In total, we collected around 300 minutes of treadmill workout data to evaluate ViBRUN. Unless stated otherwise, 70% of each subject’s data was randomly selected for training, with the remaining data equally split between validation and testing (15% each), ensuring that data from all speeds was included in each dataset. We utilized the model that demonstrated the highest performance on the validation set to evaluate its performance on the testing set.

6.2 Evaluation Metrics

We use the following metrics to evaluate the reconstruction performance of ViBRUN:

(1) **Cosine Similarity**, a widely used metric for time series data, is used to assess the accuracy of PPD reconstruction. It is calculated as:

$$\text{Cosine Similarity} = \frac{g \cdot e}{\|g\| \|e\|},$$

where g represents the ground truth PPD time series and e represents the estimated PPD.

(2) **Euclidean Distance** is used for COP reconstruction, as COP consists of x and y coordinates. The Euclidean Distance is computed as:

$$\text{Euclidean Distance} = \sqrt{(g_x - e_x)^2 + (g_y - e_y)^2},$$

where g_x and g_y are the ground truth coordinates, and e_x and e_y are the estimated coordinates.

(3) **Step Error** is employed to evaluate cadence estimation by calculating the difference in steps per minute. It is defined as:

$$\text{Step Error} = |g - e|,$$

where g is the ground truth cadence, and e is the estimated cadence.

(4) **Relative Error (RE)** is also used to evaluate the performance of cadence estimation, which is calculated as:

$$\text{RE} = \frac{|g - e|}{g},$$

providing a normalized measure of the error relative to the ground truth cadence.

(5) **Mean Absolute Error (MAE)** is used to evaluate GCT and ST estimation, capturing the discrepancy in seconds. It is calculated as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |g_i - e_i|,$$

where g_i is the ground truth and e_i is the estimated time for each test sample.

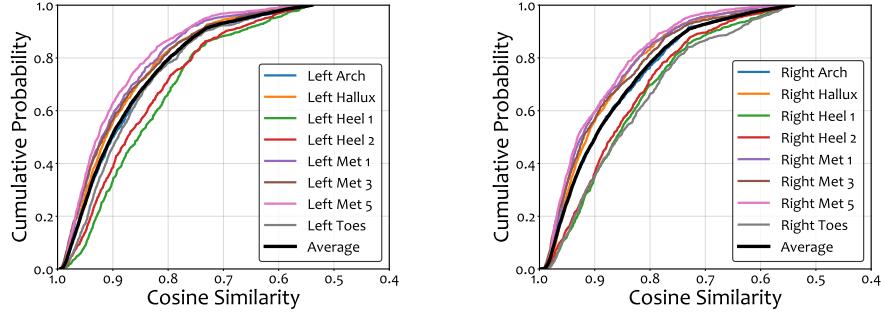


Fig. 10. CDF of PPD for 8 channels of the left and right foot and averaged cosine similarity.

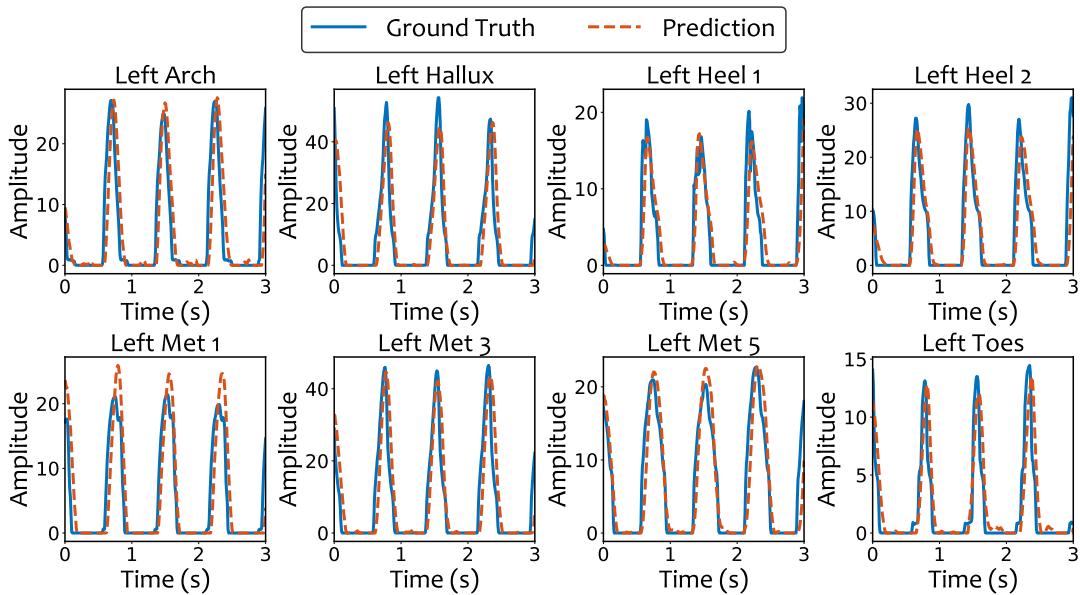


Fig. 11. PPD visualization for 8 channels of the left foot.

(6) **Root Mean Squared Error (RMSE)** is also used to evaluate COP and PPD estimation. RMSE penalizes larger errors more heavily than MAE and is calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (g_i - e_i)^2},$$

where g_i is the ground truth and e_i is the estimated time for each test sample.

6.3 System Performance

In this section, we will evaluate the performance of the VIBRUN system across all the running metrics.

6.3.1 PPD Reconstruction. Fig. 10 illustrates the Cumulative Distribution Function (CDF) of the 8-channel PPD reconstruction for both feet. We observe that across all channels for both feet, approximately 80% of the 3-second reconstructed PPD time series achieves over 80% cosine similarity with the ground truth PPD time series. Also, the average RMSE for PPD is 6.004 while the average cosine similarity is 0.879. The promising results demonstrate ViBRUN’s ability to accurately reconstruct the PPD. The average cosine similarity for the right foot ranges from 83.36% for Right Heel 1 to 89.73% for Right Met 1 with an average of 88.19%, while the left foot ranges from 83.37% for Left Heel 1 to 89.94% for Left Met 5, with an average of 87.77% across all channels for all participants. These high similarity values indicate strong alignment between the reconstructed and ground truth data across all foot regions. Fig. 11 provides an example of the reconstructed and ground truth PPD for a 3-second series across the 8 channels of the left foot. It illustrates that ViBRUN effectively captures the periodic pattern of PPD, with the reconstructed PPD highly correlated and strictly synchronized with the ground truth.

6.3.2 COP Reconstruction. The COP Y-coordinates range from 0 to 240 mm, while the X-coordinates range from 0 to 80 mm across all the datasets. Fig. 12 (a) illustrates an example of the reconstructed COP time series alongside the ground truth. Similar to the PPD reconstruction, the reconstructed COP closely mirrors the pattern of the ground truth, demonstrating high similarity. The average Euclidean Distance for the left foot across all subjects is 13.16 mm, and 13.45 mm for the right foot. Additionally, Fig. 12 (b) shows that around 80% of the COP reconstruction errors are below 20 mm for both feet, demonstrating ViBRUN’s capability to reconstruct the COP with low error. The standard deviations of 20.45 mm for the left foot and 19.68 mm for the right foot further emphasize the robustness of ViBRUN across diverse running conditions and subject variations. Also, the average RMSE for COP is 39.601 while the average cosine similarity is 0.921. This high level of accuracy provides valuable insights into the runner’s biomechanics, which can contribute to improved performance assessments and more effective injury prevention strategies.

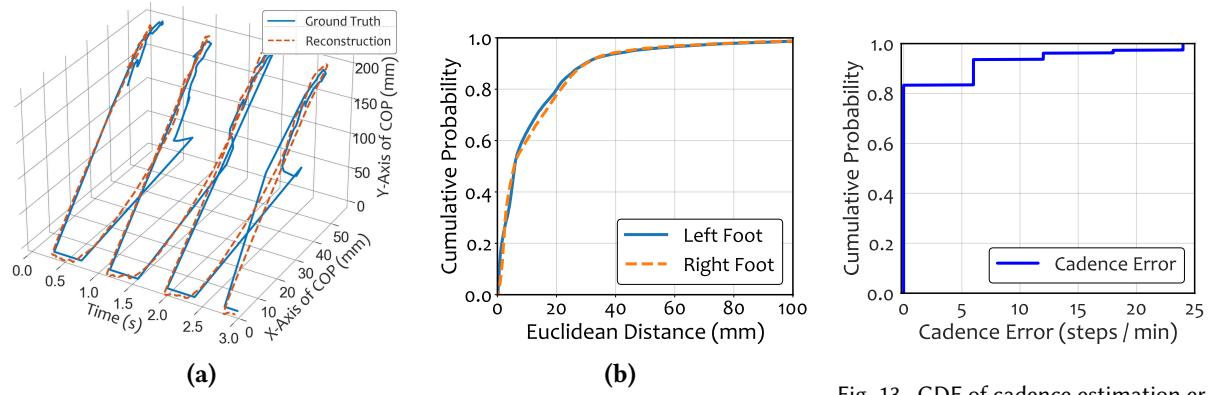


Fig. 12. Reconstruction visualizations for COP (a) and CDF of COP estimation errors (b).

Fig. 13. CDF of cadence estimation errors.

6.3.3 Cadence Estimation. We assess the performance of cadence estimation using the CDF across all subjects and speeds in Fig. 13. Approximately 83.46% of the cadence estimates achieve 0 step error in steps per minute, while around 93.72% of the estimates have a low step error within 8 steps per minute. The step error across all test samples for the cadence is 1.75 steps per minute, with a RE of 2.37%. These promising results demonstrate that ViBRUN can accurately measure the runner’s cadence, highlighting the effectiveness of our post-signal analysis algorithm.

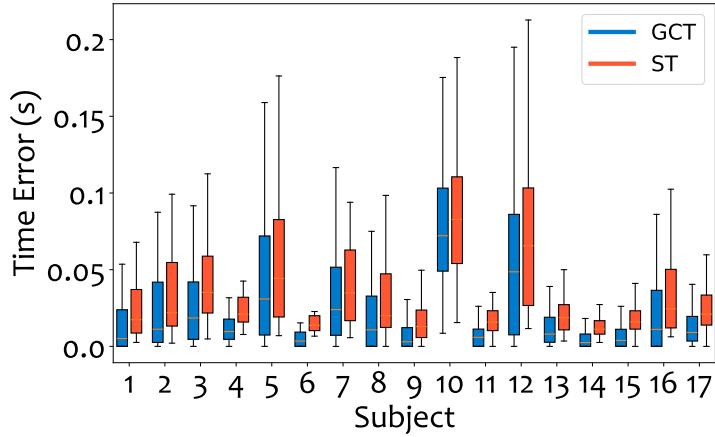


Fig. 14. GCT and ST errors for each subject.

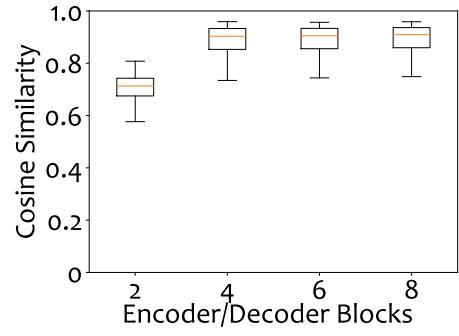


Fig. 15. Impact of the number of encoder/decoder layers on the PPD reconstruction.

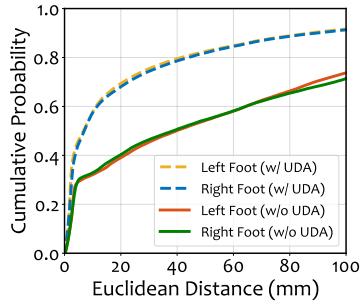


Fig. 16. CDF of Euclidean Distance for COP reconstruction after domain adaptation.

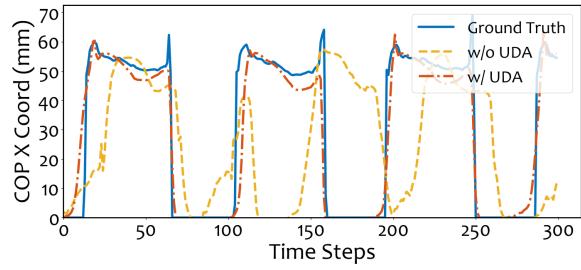


Fig. 17. Visualization of COP reconstruction comparing w/o UDA and w/ UDA models.

6.3.4 GCT & ST Estimation. Fig. 14 shows the GCT and ST estimation performance across all subjects. For most subjects, the average GCT MAE across all steps is lower than 40 ms, with subject 5 reaches the lowest MAE of 8 ms, while subject 6 has the highest seconds. The mean MAE across all subjects for GCT is 28 ms. As for ST, we observe all subjects reach a low MAE of less than 50 ms, with subject 5 having the best performance of an MAE of 11 ms, and in the worst case, the error is still only 31 ms. The average ST MAE across all subjects is 45 ms. These results demonstrate the effectiveness of VIBRUN in GCT and ST estimation. This consistency across different individuals underscores the robustness and reliability of our system.

6.3.5 Performance with Unsupervised Domain Adaptation. We evaluate the effectiveness of our domain adaptation approach under *unsupervised* settings, comparing model performance before and after adaptation. In our setup, we designate one subject as the target domain, using only their unlabeled data, while the remaining subjects' data (with labels) are used to train the model in the source domain. We observe consistent improvements in

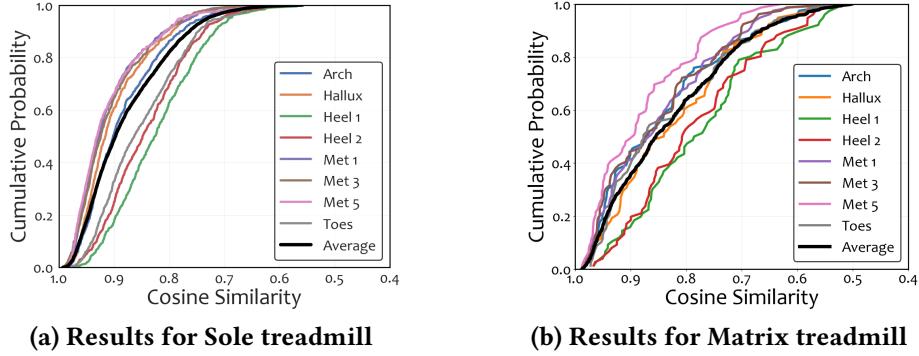


Fig. 18. Impact of different treadmills on the PPD reconstruction.

reconstructing COP and pressure signals, as well as in estimating derived gait metrics. In the unsupervised setting, where no ground truth labels are available for the target user during training, our model achieves a notable reduction in MAE for COP reconstruction, from 64.71 mm to 28.70 mm. Moreover, the CDF analysis of the Euclidean Distance (Fig. 16) shows that the model with domain adaptation achieves a distribution with significantly lower errors for both left and right feet compared with the model without domain adaptation for target user, also, a visual comparison between reconstructed COP trajectories before and after adaptation (Fig. 17) further highlights the enhanced precision enabled by unsupervised domain adaptation. For PPD reconstruction, the average RMSE is reduced from 21.55 to 13.10, and the MAE is dropped from 12.21 to 8.370. Additionally, domain adaptation improves the cosine similarity from 0.378 to 0.760. For gait metrics, the MAE for GCT, ST and cadence estimation are reduced from 101.15 ms to 59.69 ms, 133.07 ms to 77.93 ms, and 29.22 spm to 8.80 spm, respectively. Moreover, we evaluated the UDA performance using the leave-one-subject-out approach across 15 subjects, observing improvements in COP reconstruction error ranging from 22.43 mm to 35.11 mm, with an average improvement of 38 mm, demonstrating the consistency and effectiveness of our UDA approach. These results demonstrate the effectiveness of our domain adaptation strategy. These results demonstrate the effectiveness of our domain adaptation strategy.

6.4 Ablation Studies

Number of Encoder/Decoder Layers: We evaluate the impact of different numbers of layers in both the encoder and decoder on the model’s overall performance. As shown in Fig. 15, the average cosine similarity is 0.878 when the encoder/decoder blocks are 4. This increases slightly to 0.882 when the blocks are 6, indicating that there is no significant impact on performance. Since the trade-off between performance and computational efficiency is a crucial consideration for real-time deployment, we choose 4 blocks for the encoder/decoder to reduce the computational cost, ensuring that the system maintains an optimal balance between accuracy and processing speed.

Environmental Robustness: We evaluate the robustness of the model against different environments by testing its consistency across different spatial settings or configurations shown in Fig. 9. In Fig. 18, we reconstructed PPD by averaging left and right foot pressure for the same subject. Fig. 18 (a) corresponds to the scenario (a) in Fig. 9 while Fig. 18 (b) corresponds to the scenario (b) in Fig. 9. Both scenarios demonstrate satisfactory performance, with approximately 80% of the cosine similarities exceeding 0.8 for the Sole F63 treadmill, while 73% of the cosine similarities exceed 0.8 for the Matrix T5X treadmill. The relatively lower similarity on the Matrix T5X is primarily due to its structural design, which includes a more rigid frame and a thicker belt, resulting in smaller vibrations

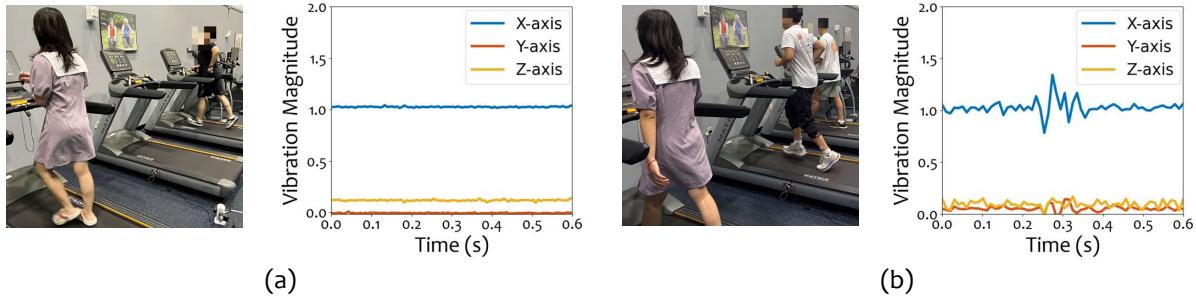


Fig. 19. Impact of nearby treadmill running on the captured vibration signals on VIBRUN (the system is mounted on the treadmill in the middle).

Table 2. Inference time with various resampled dimensions.

Resampling Dimension	300	200	100
Inference Time (s)	3.91	2.79	1.63

being transmitted to the system’s sensors. These design characteristics reduce the sensitivity of vibration-based measurements, leading to less accurate data. Despite this, the overall robustness of our VIBRUN system remains evident across different spatial settings, as the performance stays within an acceptable range.

Temporal Stability: We analyze the temporal stability of the model by assessing its performance on testing data collected approximately one month and six months apart from the same subject. The results show that the average GCT difference is only 4 ms for the one-month gap (26 ms on the same day and 30 ms one month later) and 5 ms for the six-month gap (31 ms). There is no significant performance change, highlighting the consistency and robustness of VIBRUN’s predictive capabilities over time. These results prove the temporal stability of VIBRUN, making it suitable for continuous monitoring and personalized feedback in training programs over long durations.

Impact of Nearby Treadmill Vibrations: Impact of Nearby Treadmill Vibrations: We evaluated system robustness to external noise by placing the sensor on a treadmill adjacent to active machines. As shown in Fig. 19 (a), external vibrations produced low-amplitude signals compared to those in Fig. 19 (b), indicating minimal interference. These results demonstrate that nearby treadmill activity does not significantly impact reconstruction accuracy, confirming VibRun’s reliability in shared gym or multi-treadmill environments.

Real-time Inference Study: To ensure smooth real-time operation, inference time must remain below the 3-second data segmentation window. We evaluated VibRun’s adaptability to various computational capacities by resampling input sequences to 300, 200, and 100 samples while maintaining a fixed 3-second window via adjusted sampling rates. As shown in Table 2, using 100 samples achieves real-time performance on the Raspberry Pi (1.63 s), balancing processing speed and running metric accuracy.

7 RELATED WORK

In this section, we review existing running performance monitoring systems that utilize either wearable or contactless sensors.

7.1 Wearable sensor based

Moov Now™ [58], MilestonePod [59], RunScribe™ [60], and Zoi [61] are commercially available wearable devices for running gait analysis. However, these devices are limited in their ability to provide comprehensive gait analysis, as they typically only offer a narrow range of metrics, such as cadence, and fail to deliver a complete assessment of running gait [62, 63]. Instead of relying on market-available devices, recent research has shifted towards utilizing IMU sensors [8, 64, 65], pressure sensors, and smart insoles [66] to determine running gait parameters more accurately. IMU sensors affixed to shoes can estimate foot kinematics, such as foot count and foot strike [12], heading direction, speed [14], and stride length [15]. Additionally, IMU sensors can be attached to the shank, thigh, and foot [67, 68] to measure angular motions and calculate stride lengths, heel strikes, and toe-off events. To mitigate the impact of angular motions, IMUs have also been placed on the pelvis to detect foot contacts [69]. Systems such as RunBuddy [70] and PaceGuard [71] employ smartphone accelerometers and auditory feedback to measure running rhythm and cadence. These smartphones are attached to the body during running. The PhysioTreadmill [72] uses an intelligent band with an accelerometer and gyroscope to capture physiological data in VR.

While wearable devices have successfully predicted multiple metrics for running gait estimation, they are often intrusive due to body contact. Seuter *et al.* investigate how interacting with wearable devices, such as smartphones, smartwatches, and smartglasses, affects running movement, and the study shows that these interactions introduce interference, highlighting the drawbacks of wearable devices during physical activities [10]. Moreover, these research efforts typically focus on specific body parts to predict particular metrics and do not offer a comprehensive model that estimates all metrics. In contrast, our proposed system can predict five running metrics, is cost-effective, and employs a contactless approach for an improved user experience.

7.2 Contactless sensor based

Contactless sensing provides an unobtrusive solution for monitoring running gait. Various contactless methods have been employed for gait analysis (walking) and vital signs monitoring, including acoustic-based methods [19, 20], video analysis [22, 23], 3D motion capture using multiple cameras [24], and specialized sensors such as radar [25, 26] and lidar [27]. However, methods specifically for monitoring and analyzing running gait using contactless sensors are limited [73]. Although 3D motion capture is often regarded as the gold standard for contactless running gait monitoring, these systems typically require expensive multi-camera setups or specialized hardware and software [28, 29]. Existing video-based methods for running gait analysis rely on pose estimation models [74, 75] or markers attached to the runner's body [76] to track anatomical landmarks, such as joint locations. While effective, these approaches often collect sensitive visual data that contain identifiable personal information, raising privacy concerns. A recent study utilized acoustic signals collected from a smartphone placed on a treadmill to reconstruct metrics such as cadence and ground contact time [21]. However, the range of metrics covered is limited, and audio signals are susceptible to distortion from background noise. In contrast to existing contactless methods, our system is significantly more cost-effective, privacy-preserving, and provides a comprehensive monitoring solution. It can track a wide range of running metrics, including cadence, ground contact time, center of pressure, and pressure distribution, thereby offering a more detailed and thorough analysis.

8 DISCUSSION

8.1 Comparison with Baselines for Performance Insight

To better understand the performance of ViBRUN, we compare our results against existing systems, including both commercial devices and research prototypes. Although these comparisons are not entirely fair due to differences in hardware, data collection setups, and target activities, they still provide valuable insights into how well our system performs in practice. For cadence estimation, Apple Watch has a reported error of about 10.69 spm, while

foot-mounted sensors like RunScribe achieve much lower errors, around 1.54 spm, and a recent acoustic sensing system achieves 3.74 spm [77]. In comparison, VIBRUN achieves a MAE of 1.75 spm, which is very close to the most accurate commercial devices while relying on a simple, unobtrusive IMU sensor.

For GCT, the RunScribe system reports an error of 22.15 ms, and the acoustic system named MTL-MLP reports 23.86 ms [77]. The error estimate when the IMU sensor is attached to the back, arm, and foot/shoes is 8 ms, 49 ms, and 14ms, respectively [78]. VIBRUN achieves an MAE of 28 ms, which remains within a comparable range and demonstrates its ability to provide reliable results without requiring on-body sensors. MAE for ST with xGB and Apple Watch Data is observed as around 70 ms [79]. In contrast, our model achieves an MAE of around 44 ms for ST estimation, indicating strong performance.

For COP reconstruction, Chan *et al.* [80] use a 99-sensor insole system and achieves an average error of 8.52 mm during walking. In comparison, VIBRUN focuses on running, a more intense and complex motion, using only a single accelerometer. Despite the simpler sensor setup and the more challenging task, we achieve a COP error of 14.84 mm, which is reasonable given the context. In general, these results suggest that VIBRUN can offer reliable performance in multiple metrics, even compared to more complex or commercial systems. Although exact comparisons are not always possible, our results demonstrate that a simple, low-cost, and unobtrusive setup can still provide useful and accurate insights for running analysis.

8.2 Limitations and Future Work

While VibRun offers promising performance in unobtrusive running gait monitoring, we acknowledge the system has some limitations that we intend to address in future work.

Experimental Scope and Real-World Robustness: Since our current evaluation is conducted in controlled indoor treadmill environments, real-world deployment may face additional challenges. Practical running scenarios often involve factors such as treadmill incline changes, motion artifacts (e.g., drinking water, sudden speed adjustments), and variations in sensor placement. These dynamics are not currently represented in our training data. For example, while mounting the IMU sensor on the middle side of the treadmill yields the most consistent vibration signals—due to the direct impact of foot strikes—alternative placements at the front or back may reduce signal clarity or introduce phase shifts caused by treadmill belt elasticity. To improve robustness and generalizability, future work will systematically incorporate such factors by expanding the dataset to cover diverse treadmill types, incline configurations, usage scenarios, and sensor placements.

Cross-user Generalization and Personalization: Another important limitation of our system is the variability in vibration signals across different individuals because of personal gait patterns and biomechanics. This inter-subject variability poses challenges for model generalization. To address this, we employ UDA techniques, which facilitate model adaptation to new subjects without requiring labeled data. While our approach shows promise, the results for unseen subjects still need to be improved. One promising direction is training on larger and more diverse participant datasets. Some practical systems, like Egolm [81], HMD² [82], use the Nymeria dataset from Meta [83] recordings from 264 participants, highlighting the importance of population-scale data collection for improving model robustness and generalization in real-world applications. Inspired by such efforts, future work will focus on expanding our dataset to encompass a broader range of users and movement scenarios, enhancing VibRun’s adaptability to individual differences. In addition, unsupervised meta-learning [84] could help address individual gait variability with minimal new data and may also serve as a promising direction for future work.

9 CONCLUSION

In this paper, we introduce VIBRUN, a novel unobtrusive gait analysis system designed for indoor treadmill running. VIBRUN aims to enhance running efficiency and prevent potential injuries by providing real-time tracking of five critical running metrics: cadence, ground contact time, stride time, center of pressure, and plantar

pressure distribution. Unlike traditional methods, ViBRUN can be easily deployed on treadmills, eliminating the need for runners to wear any devices, and ensuring minimal obtrusiveness while remaining cost-effective and computationally efficient. Utilizing a multi-task transformer model, comprehensive evaluations demonstrate that ViBRUN accurately monitors these five metrics with low mean absolute error and high cosine similarity. The system's adaptability for real-time applications and resilience against external noise further emphasize its potential for real-world use. We envision that ViBRUN will pave the way for new research opportunities in virtual sports events and gamified fitness, enhancing user engagement and reducing injury risks in online training sessions.

10 ACKNOWLEDGMENTS

We sincerely thank the anonymous reviewers for their thoughtful and constructive feedback. This work was supported in part by NSF CNS-2114161, ECCS2132106, CBET-2130643, and CNS-2403529.

REFERENCES

- [1] Number of running and jogging participants in the united states from 2010 to 2023. <https://www.statista.com/statistics/190303/running-participants-in-the-us-since-2006/>, 2024.
- [2] Treadmill statistics 2024 by types, components, mechanisms. <https://www.news.market.us/treadmill-statistics/>, 2024.
- [3] Boris Smus and Vassilis Kostakos. Running gestures: hands-free interaction during physical activity. In *Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing-Adjunct*, pages 433–434, 2010.
- [4] Zwift. Your pace your place the free app for running and h your community., 2023.
- [5] Kobe Helsen, Inge Derom, Joris Corthouts, Veerle De Bosscher, Annick Willem, and Jeroen Scheerder. Participatory sport events in times of covid-19: Analysing the (virtual) sport behaviour of event participants. *European Sport Management Quarterly*, 22(1):35–54, 2022.
- [6] Erisher Woyo and Costantine Nyamandi. Application of virtual reality technologies in the comrades' marathon as a response to covid-19 pandemic. *Development Southern Africa*, 39(1):20–34, 2022.
- [7] Zwift. Popular running events on zwift. <https://www.zwift.com/eu-es/news/21470-popular-running-events-on-zwift>, 2024.
- [8] Matthias Seuter, Lucien Opitz, Gernot Bauer, and David Hochmann. Live-feedback from the imu: Animated 3d visualization for everyday-exercising. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 904–907, 2016.
- [9] Rúben Gouveia, Fábio Pereira, Evangelos Karapanos, Sean A Munson, and Marc Hassenzahl. Exploring the design space of glanceable feedback for physical activity trackers. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*, pages 144–155, 2016.
- [10] Matthias Seuter, Max Pfeiffer, Gernot Bauer, Karen Zentgraf, and Christian Kray. Running with technology: Evaluating the impact of interacting with wearable devices on running movement. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):1–17, 2017.
- [11] Christine F Martindale, Markus Wirth, Stefan Schneegass, Markus Zrenner, Benjamin H Groh, Peter Blank, Dominik Schulzhaus, Thomas Kautz, and Bjoern M Eskofier. Workshop on wearables for sports. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 877–880, 2016.
- [12] GP Bailey and R Harle. Assessment of foot kinematics during steady state running using a foot-mounted imu. *Procedia Engineering*, 72:32–37, 2014.
- [13] Mahmoud Hassan, Florian Daiber, Frederik Wiehr, Felix Kosmalla, and Antonio Krüger. Footstriker: An ems-based foot strike assistant for running. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(1):1–18, 2017.
- [14] Yi-Cheng Huang, Yu-Rui Chen, Hung-Yi Wu, and Yu-Jui Huang. Wearable sensor for measurement of gait walking and running motion. *Sensors & Materials*, 31, 2019.
- [15] Lauren C Benson, Christian A Clermont, and Reed Ferber. New considerations for collecting biomechanical data using wearable sensors: the effect of different running environments. *Frontiers in Bioengineering and Biotechnology*, 8:86, 2020.
- [16] Tekscan. F-scan64. <https://www.tekscan.com/products-solutions/systems/f-scan64>, 2023.
- [17] Ultium insoles. <https://www.noraxon.com/our-products/insole-smartlead>, 2024.
- [18] Ha Le, Rithika Lakshminarayanan, Jixin Li, Varun Mishra, and Stephen Intille. Collecting self-reported physical activity and posture data using audio-based ecological momentary assessment. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 8(3):1–35, 2024.
- [19] Chao Cai, Henglin Pu, Peng Wang, Zhe Chen, and Jun Luo. We hear your pace: Passive acoustic localization of multiple walking persons. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(2):1–24, 2021.

- [20] M Umair Bin Altaf, Taras Butko, and Biing-Hwang Juang. Acoustic gaits: Gait analysis with footstep sounds. *IEEE Transactions on Biomedical Engineering*, 62(8):2001–2011, 2015.
- [21] Ziyi Xuan, Ming Liu, Jingping Nie, Minghui Zhao, Stephen Xia, and Xiaofan Jiang. Canrun: Non-contact, acoustic-based cadence estimation on treadmills using smartphones. In *Proceedings of Cyber-Physical Systems and Internet of Things Week 2023*, pages 272–277. 2023.
- [22] Bochao Zou, Zizheng Guo, Xiaocheng Hu, and Huimin Ma. Rhythmmamba: Fast remote physiological measurement with arbitrary length videos. *arXiv preprint arXiv:2404.06483*, 2024.
- [23] Sandy Ardianto and Hsueh-Ming Hang. Multi-view and multi-modal action recognition with learned fusion. In *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 1601–1604. IEEE, 2018.
- [24] Justin Amadeus Albert, Victor Owolabi, Arnd Gebel, Clemens Markus Brahms, Urs Granacher, and Bert Arnrich. Evaluation of the pose tracking performance of the azure kinect and kinect v2 for gait analysis in comparison with a gold standard: A pilot study. *Sensors*, 20(18):5104, 2020.
- [25] Tianyue Zheng, Zhe Chen, Shujie Zhang, Chao Cai, and Jun Luo. More-fi: Motion-robust and fine-grained respiration monitoring via deep-learning uwb radar. In *Proceedings of the 19th ACM conference on embedded networked sensor systems*, pages 111–124, 2021.
- [26] Zhe Chen, Tianyue Zheng, Chao Cai, and Jun Luo. Movi-fi: Motion-robust vital signs waveform recovery via deep interpreted if sensing. In *Proceedings of the 27th annual international conference on mobile computing and networking*, pages 392–405, 2021.
- [27] Angela Botros, Nathan Gyger, Narayan Schütz, Michael Single, Tobias Nef, and Stephan M Gerber. Contactless gait assessment in home-like environments. *Sensors*, 21(18):6205, 2021.
- [28] Angkoon Phinyomark, Sean Osis, Blayne A Hettinga, and Reed Ferber. Kinematic gait patterns in healthy runners: A hierarchical cluster analysis. *Journal of biomechanics*, 48(14):3897–3904, 2015.
- [29] Michael B Pohl, Chandra Lloyd, and Reed Ferber. Can the reliability of three-dimensional running kinematics be improved using functional joint methodology? *Gait & Posture*, 32(4):559–563, 2010.
- [30] What is running cadence? it might be the key to running longer. <https://www.healthline.com/health/fitness/what-is-cadence-in-running#why-its-important>, 2023.
- [31] Running cadence explained: What is a good running cadence? <https://marathonhandbook.com/running-cadence/>, 2023.
- [32] Thomas Dos’ Santos, Christopher Thomas, Paul Comfort, and Paul A Jones. The effect of angle and velocity on change of direction biomechanics: An angle-velocity trade-off. *Sports medicine*, 48:2235–2253, 2018.
- [33] James R Andrews, Gary L Harrelson, and Kevin E Wilk. *Physical Rehabilitation of the Injured Athlete: Expert Consult-Online and Print*. Elsevier Health Sciences, 2012.
- [34] 3 steps to a better stride. <https://www.trailrunnermag.com/training/trail-tips-training/3-steps-to-a-better-stride/>, 2022.
- [35] Amy G Schubert, Jenny Kempf, and Bryan C Heiderscheit. Influence of stride frequency and length on running mechanics: a systematic review. *Sports health*, 6(3):210–217, 2014.
- [36] Seyed Hamed Mousavi, Laurens van Kouwenhove, Reza Rajabi, Johannes Zwerver, and Juha M Hilmans. The effect of changing mediolateral center of pressure on rearfoot eversion during treadmill running. *Gait & Posture*, 83:201–209, 2021.
- [37] Thomas J Hagedorn, Alyssa B Dufour, Yvonne M Golightly, Jody L Riskowski, Howard J Hillstrom, Virginia A Casey, and Marian T Hannan. Factors affecting center of pressure in older adults: the framingham foot study. *Journal of foot and ankle research*, 6(1):1–5, 2013.
- [38] Dong-ping Wan, Huan-li Bao, Jian-peng Wang, Jie Wei, Jian-bing Ma, Shu-xin Yao, and Chao Xu. Plantar pressure distribution and posture balance during walking in individuals with unilateral chronic ankle instability: An observational study. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, 29:e940252–1, 2023.
- [39] Daniel J Inman and Ramesh Chandra Singh. *Engineering vibration*, volume 3. Prentice Hall Englewood Cliffs, NJ, 1994.
- [40] Amit Pandey and Achin Jain. Comparative analysis of knn algorithm using various normalization techniques. *International Journal of Computer Network and Information Security*, 10(11):36, 2017.
- [41] Md Manjurul Ahsan, MA Parvez Mahmud, Pritom Kumar Saha, Kishor Datta Gupta, and Zahed Siddique. Effect of data scaling methods on machine learning algorithms and model performance. *Technologies*, 9(3):52, 2021.
- [42] myomuscle 3.6 user guide. <https://www.noraxon.com/download/myomuscle-3-6-user-guide/>, 2024.
- [43] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [44] Qingsong Wen, Tian Zhou, Chaoli Zhang, Weiqi Chen, Ziqing Ma, Junchi Yan, and Liang Sun. Transformers in time series: A survey. *arXiv preprint arXiv:2202.07125*, 2022.
- [45] Deblina Bhattacharjee, Tong Zhang, Sabine Süssstrunk, and Mathieu Salzmann. Mult: An end-to-end multitask learning transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12031–12041, 2022.
- [46] Garrett Wilson and Diane J Cook. A survey of unsupervised deep domain adaptation. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 11(5):1–46, 2020.
- [47] Ground contact time. <https://sporttracks.mobi/blog/what-is-ground-contact-time-in-running>, 2017.

- [48] Catrine Tudor-Locke, Ho Han, Elroy J Aguiar, Tiago V Barreira, John M Schuna Jr, Minsoo Kang, and David A Rowe. How fast is fast enough? walking cadence (steps/min) as a practical estimate of intensity in adults: a narrative review. *British journal of sports medicine*, 52(12):776–788, 2018.
- [49] Liam P Pellerine, Jennifer L Petterson, Madeline E Shulgulam, Peter J Johansson, Pasan Hettiarachchi, Derek S Kimmerly, Ryan J Frayne, and Myles W O’Brien. Step length, but not stepping cadence, strongly predicts physical activity intensity during jogging and running. *Measurement in Physical Education and Exercise Science*, 27(4):352–361, 2023.
- [50] Meng-Shiuan Pan and Hsueh-Wei Lin. A step counting algorithm for smartphone users: Design and implementation. *IEEE Sensors Journal*, 15(4):2296–2305, 2014.
- [51] Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [52] Matrix t5x treadmill sell sheet – matrix fitness. https://content.johnsonfit.com/inc/uploaded_media/bf0bd876475c30601b23191f0e7d3973/sell_sheet/afed41c1418fba221c72e5b70860bb02.pdf, 2025.
- [53] Spirit xt685 treadmill review – treadmill doctor. <https://www.treadmilldoctor.com/spirit-xt685-treadmill-review>, 2024.
- [54] Xt685 light commercial treadmill – the fitness market. https://www.thefitnessmarket.com/Spirit-XT685-Light-Commercial-Treadmill_p_17.html, 2025.
- [55] Best cushioned treadmills for shock-absorbing exercise – sole fitness blog. <https://www.soletreadmills.com/blogs/news/best-cushioned-treadmills-for-shock-absorbing-exercise-2024>, 2024.
- [56] Sole f63 treadmill belt – treadmillrunningbelts.com.au. <https://treadmillrunningbelts.com.au/product/sole-f63/>, 2025.
- [57] Sole f63 treadmill – sole fitness canada. <https://solefitness.ca/f63-treadmill>, 2025.
- [58] Train right with moov now™. <https://welcome.moov.cc/>, 2024.
- [59] Hands-on with the milestone pod. <https://www.cnet.com/roadshow/pictures/milestonepod-pictures/>, 2024.
- [60] Runscribe gait lab – gait analysis system. <https://runscribe.com/>, 2024.
- [61] Zoi - smart wearable to find your running beat. <https://www.indiegogo.com/projects/zoi-smart-wearable-to-find-your-running-beat/>, 2024.
- [62] B. Pariot de Fontenay, J. S. Roy, B. Dubois, L. Bouyer, and J. F. Esculier. Validating commercial wearable sensors for running gait parameters estimation. *IEEE Sensors Journal*, 20(14):7783–7791, 2020.
- [63] Emilio J Ruiz-Malagón, Felipe García-Pinillos, Alejandro Molina-Molina, Víctor M Soto-Hermoso, and Santiago A Ruiz-Alias. Runscribe sacral gait lab™ validation for measuring pelvic kinematics during human locomotion at different speeds. *Sensors*, 23(5):2604, 2023.
- [64] E Parikesit, TLR Mengko, and H Zakaria. Wearable gait measurement system based on accelerometer and pressure sensor. In *2011 2nd International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering*, pages 395–398. IEEE, 2011.
- [65] Rajesh Kumar, Can Isik, and Vir V Phoha. Treadmill assisted gait spoofing (tags) an emerging threat to wearable sensor-based gait authentication. *Digital Threats: Research and Practice*, 2(3):1–17, 2021.
- [66] Kyoungchul Kong and Masayoshi Tomizuka. A gait monitoring system based on air pressure sensors embedded in a shoe. *IEEE/ASME Transactions on mechatronics*, 14(3):358–370, 2009.
- [67] Shuzhi Yang, Chris Mohr, and Qingguo Li. Ambulatory running speed estimation using an inertial sensor. *Gait & posture*, 34(4):462–466, 2011.
- [68] A.U. Alahakone, S.M.N.A. Senanayake, and C.M. Senanayake. Smart wearable device for real time gait event detection during running. In *2010 IEEE Asia Pacific Conference on Circuits and Systems*, pages 612–615, 2010.
- [69] Jasper Reenalda, Marit A Zandbergen, Jelle HD Harbers, Max R Paquette, and Clare E Milner. Detection of foot contact in treadmill running with inertial and optical measurement systems. *Journal of biomechanics*, 121:110419, 2021.
- [70] Tian Hao, Guoliang Xing, and Gang Zhou. Runbuddy: a smartphone system for running rhythm monitoring. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, pages 133–144, 2015.
- [71] Jutta Fortmann, Martin Pielot, Marco Mitteldorf, Martin Büscher, Stefan Trienen, and Susanne Boll. Paceguard: improving running cadence by real-time auditory feedback. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services companion*, pages 5–10, 2012.
- [72] Shaolong Liu, Xingce Wang, Zhongke Wu, and Ying He. Physiotreadmill: an auto-controlled treadmill featuring physiological-data-driven visual/audio feedback. In *2020 International Conference on Cyberworlds (CW)*, pages 219–226. IEEE, 2020.
- [73] Praneeth Susarla, Anirban Mukherjee, Manuel Lage Cañellas, Constantino Álvarez Casado, Xiaoting Wu, Olli Silvén, Dinesh Babu Jayagopi, Miguel Bordallo López, et al. Non-contact multimodal indoor human monitoring systems: A survey. *Information Fusion*, 110:102457, 2024.
- [74] Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, and Matthias Grundmann. Blazepose: On-device real-time body pose tracking. *arXiv preprint arXiv:2006.10204*, 2020.
- [75] Fraser Young, Rachel Mason, Rosie Morris, Samuel Stuart, and Alan Godfrey. Internet-of-things-enabled markerless running gait assessment from a single smartphone camera. *Sensors*, 23(2):696, 2023.

- [76] Seyed Hamed Mousavi, Juha M Hijmans, Forough Moeini, Reza Rajabi, Reed Ferber, Henk van der Worp, and Johannes Zwerver. Validity and reliability of a smartphone motion analysis app for lower limb kinematics during treadmill running. *Physical Therapy in Sport*, 43:27–35, 2020.
- [77] Jingping Nie, Yuang Fan, Ziyi Xuan, Matthias Preindl, and Xiaofan Jiang. Real-time non-contact estimation of running metrics on treadmills using smartphones. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*, pages 1644–1646, 2024.
- [78] Leticia González, Antonio M López, Diego Álvarez, and Juan C Álvarez. Estimation of ground contact time with inertial sensors from the upper arm and the upper back. *Sensors*, 23(5):2523, 2023.
- [79] Christopher A Bailey, Alexandre Mir-Orefice, Thomas K Uchida, Julie Nantel, and Ryan B Graham. Smartwatch-based prediction of single-stride and stride-to-stride gait outcomes using regression-based machine learning. *Annals of biomedical engineering*, 51(11):2504–2517, 2023.
- [80] Hsiao-Lung Chan, Jing-Rong Liang, Ya-Ju Chang, Rou-Shayn Chen, Cheng-Chung Kuo, Wen-Yen Hsu, and Meng-Tsan Tsai. Enhancing plantar pressure distribution reconstruction with conditional generative adversarial networks from multi-region foot pressure sensing. *Biomedical Signal Processing and Control*, 100:107187, 2025.
- [81] Fangzhou Hong, Vladimir Guzov, Hyo Jin Kim, Yuting Ye, Richard Newcombe, Ziwei Liu, and Lingni Ma. Egolm: Multi-modal language model of egocentric motions. *arXiv preprint arXiv:2409.18127*, 2024.
- [82] Vladimir Guzov, Yifeng Jiang, Fangzhou Hong, Gerard Pons-Moll, Richard Newcombe, C Karen Liu, Yuting Ye, and Lingni Ma. HmdΩ: Environment-aware motion generation from single egocentric head-mounted device. *arXiv preprint arXiv:2409.13426*, 2024.
- [83] Lingni Ma, Yuting Ye, Fangzhou Hong, Vladimir Guzov, Yifeng Jiang, Rowan Postyeni, Luis Pesqueira, Alexander Gamino, Vijay Baiyya, Hyo Jin Kim, et al. Nymeria: A massive collection of multimodal egocentric daily motion in the wild. In *European Conference on Computer Vision*, pages 445–465. Springer, 2024.
- [84] Kyle Hsu, Sergey Levine, and Chelsea Finn. Unsupervised learning via meta-learning. *arXiv preprint arXiv:1810.02334*, 2018.