In-home Gait Abnormality Detection through Footstep-Induced Floor Vibration Sensing and Person-Invariant Contrastive Learning

Yiwen Dong, Sung Eun Kim, Kornél Schadl, Peide Huang, Wenhao Ding, Jessica Rose, and Hae Young Noh

***Abstract*— Detecting gait abnormalities is crucial for assessing fall risks and early identification of neuromusculoskeletal disorders such as Parkinson's and stroke. Traditional assessments in gait clinics are infrequent and pose barriers, particularly for disadvantaged populations. Previous efforts have explored sensor-based approaches for in-home gait assessments, yet they face limitations such as visual obstructions (cameras), limited coverage (pressure mats), and the need for device carrying (wearables and insoles). To overcome these limitations, we introduce an in-home gait abnormality detection system using footstep-induced floor vibrations, enabling low-cost, non-intrusive, device-free gait health monitoring. The main research challenge is the high uncertainty in floor vibrations due to gait variations among people, making it challenging to develop a generalizable model for new patients. To address this, we analyze time-frequency-domain features of floor vibration data during specific gait phases and develop a feature transformation method through contrastive learning to address the between-people gait variation challenge. Our method transforms the features from vibrations to an embedding space where samples from different people stay close to each other (robust to people variation) while normal and abnormal gait samples are far apart (sensitive to gait abnormality). After that, gait abnormalities are detected by a downstream classifier after feature transformation. We evaluated our approach through a real-world walking experiment with 21 participants and achieved an 85% to 95% mean accuracy in detecting various gait abnormalities. This novel method overcomes prior limitations in in-home gait assessments, offering accessible gait abnormality detection without the need for intrusive devices or labels for new patients.**

***Index Terms*— Gait assessment, Gait disorders, Contrastive Learning, Sensors, Vibrations**

This work was supported in part by the U.S. National Science Foundation (under grant numbers NSF-CMMI-2026699).

Yiwen Dong and Hae Young Noh are with Department of Civil and Environmental Engineering, Stanford University, 473 Via Ortega, Stanford, CA 94305 USA (e-mail: ywdong@stanford.edu, noh@stanford.edu).

Sung Eun Kim, Kornél Schadl, and Jessica Rose are with both Motion & Gait Analysis Lab, Lucile Packard Children's Hospital, Menlo Park, CA 94025 USA and Department of Orthopaedic Surgery, Stanford University, Redwood City, CA 94063 USA (e-mail: cjsekim@stanford.edu, kornels@stanford.edu, Jessica.Rose@stanford.edu).

Peide Huang and Wenhao Ding are with Department of Mechanical Engineering, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213 USA (e-mail: peideh@andrew.cmu.edu, wenhaod@andrew.cmu.edu).

I. Introduction

Detecting gait abnormalities is essential for early identification of the risk of trips and falls, and progressive/rehabilitation tracking of neuromusculoskeletal disorders, such as joint impairments, Parkinson's, and stroke [1]–[4]. Gait abnormalities are typically reflected in deviations of posture from normal walking patterns and indicate underlying conditions that affect neuromusculoskeletal systems [5]. Among them, analyzing how the foot contacts and interacts with the floor during walking is a critical aspect because it helps to understand pathological conditions in gait such as foot drag and toe-walking, as well as inferring how the ground reaction forces act and are transmitted through the body [6].

Existing gait analysis relies on observational gait analysis and the use of specialized sensing systems in well-calibrated gait clinics such as motion capture systems, force plates, and Electromyography (EMG) [7]–[9]. However, they require in-person visits to gait clinics, often resulting in infrequent and discontinuous results under controlled walking settings. These systems are also expensive and require specialized installation, and professionally trained staff to operate. As a result, they also pose barriers for disadvantaged populations to access healthcare services, especially for low-income families who live in remote areas. To overcome these issues, studies have explored portable devices that are more practical for everyday use, such as video cameras, pressure mats, and wearables [10]–[17]. However, the cameras require direct line-of-sight and is sensitive to changing light conditions, and their measurements are limited to body motion, making it difficult to infer abnormal footstep force exerted onto the floor. The pressure mats provide force-related information, but they require sense deployment and thus have limited areas of coverage. The wearables enable more ubiquitous usage, but they require a person to carry devices on designated body parts (e.g., wrist, legs, ankle), which may cause discomfort and inconvenience. To overcome these limitations, prior works have leveraged footstep-induced floor vibration sensing for gait health monitoring [18]–[23], which is device-free, wide-ranging, and perceived as more privacy-friendly. While floor vibrations are sensitive to environmental disturbances, prior studies have addressed these concerns through noise filtering techniques and classification algorithms [24], [25]. For gait

health monitoring, existing studies using this approach have been successful in extracting basic parameters such as step time and step frequency [26]–[28]. However, these approaches have not looked into complex gait characteristics during initial contact, stance, and foot clearance, making them unable to detect gait abnormalities that are meaningful in clinical settings. Moreover, the existing studies have not explored the effect of various gait types, sensor layouts, or individual differences on gait health monitoring.

In this study, we model footstep-induced floor vibrations and develop gait phase-informed features from the vibration signals within three representative stages in a gait cycle to provide fine-grained, interpretable, and reliable in-home gait assessments. The physical insight of this approach is that as a person's foot contacts the floor during each stage of a gait cycle, there are significant differences in contact forces between normal and abnormal gait, such as the angle, magnitude, area, and duration. Such differences in contact forces generate distinctive vibration wave patterns on the floor. As these vibration waves propagate through the floor, we capture their patterns by installing vibration sensors on the floor surface. By analyzing the vibration signals, we infer the characteristics of the dynamic foot-floor contact forces within each stage of a gait cycle, including initial contact type, weight translation speed, and quality of foot clearance, which allows gait abnormality detection. Our approach is non-intrusive, wide-ranged, and does not require device wearing, which is suitable for continuous in-home monitoring, which increases the availability and accessibility of gait assessment services.

The main research challenge in developing our approach is the high uncertainty in floor vibrations due to the variations in walking styles among people. Prior studies found that each individual has a unique gait, as a result of people's body configurations and walking habits [29]. This affects the interaction pattern between the foot and the floor [25], [30]–[32] and is mixed with the effect of gait abnormalities, making it challenging to determine whether a certain deviation from a typical normal gait is due to that person's unique walking habits or not. To overcome these challenges, we focus on 3 critical stages within a gait cycle that are of major concern in clinical studies. They are 1) the initial contact stage, 2) the weight translation stage during the stance phase, and 3) the foot clearance stage during the swing phase. We first extract gait phase-informed features from floor vibration signals during these 3 stages to represent normal and abnormal gait characteristics. We then develop a person-invariant gait feature transformation framework based on contrastive learning, with the context of person and gait type information in order to address the challenge of gait differences among people. Conventional contrastive learning aims to extract features from unlabelled input data that are similar within the same given category and dissimilar outside of that category. To make our method person-invariant, we leverage contrastive learning to transform the vibration features into a space such that footstep samples from different gait types (normal vs. abnormal) are far apart and those from the same gait type stay close to each other regardless of the gait differences among people. After the transformation, we compare the embeddings of newly observed footstep samples with that of the normal footsteps to detect gait abnormalities, which translates the model estimations into interpretable gait health insights tailored to each person's gait. The output of our approach is visualized through a personalized risk profile during each critical stage showing the distribution of all recorded gait cycles.

The contributions of the study are that we:

1) Develop the first floor vibration-based gait abnormality detection system. Our system detects abnormal foot-floor interactions during initial contact, stance, and swing phase for efficient and non-intrusive gait health monitoring at home;
2) Model the gait phase-informed features within a gait cycle through a person-invariant feature transformation based on contrastive learning to reduce uncertainties due to the gait variations among people;
3) Evaluate the approach through a real-world walking experiment with 21 participants and 60,000+ gait cycles to demonstrate its effectiveness and robustness.

We conducted a real-world walking experiment with 21 participants. Our approach achieves an 85%, 90%, and 95% mean test accuracy in detecting gait abnormalities for each gait cycle during initial contact, stance phase, and swing phase, respectively. Overall, our approach has an average of 0.81 test F1 score in classifying four types of gait, including normal gait, gait with flatfoot/toe initial contact, gait with flexed-knee in stance phase, and gait with foot drag in swing phase, which establishes its efficacy in in-home gait assessments.

II. Related Work

In this section, we summarize the related work on gait abnormality detection, mainly from two perspectives: 1) existing sensing systems, and 2) prior work on floor vibration-based gait monitoring.

A. Existing Sensing Systems on Gait Health Monitoring

Existing sensing systems in gait health monitoring are summarized in Table II-A to include their qualitative advantages and disadvantages in real-world settings.

Based on the comparison above, gait health monitoring in non-clinical settings presents considerable challenges. Specifically, users' expectations regarding convenience, comfort, and privacy protection are notably higher, while the daily living environment is typically noisier and more complex compared to clinical settings. Despite these challenges, floor vibration sensing offers distinct advantages: it is non-invasive, covers a wide range, and is perceived as more privacy-friendly for daily use. However, this approach remains relatively under-explored, and challenges related to the floor types and between-person variations have yet to be adequately addressed. Therefore, our study and future work will focus on floor vibration sensing and endeavor to develop solutions to overcome its challenges.

TABLE I
QUALITATIVE COMPARISON OF THE EXISTING SENSING SYSTEMS FOR GAIT ANALYSIS

Sensing Method	Setting (C/NC)*	Advantages	Disadvantages
MoCap System	C	Accurate body motion	Expensive; Require professional training and calibration
EMG	C	Capture muscle activation	Limited muscle coverage
Force Plates	C/NC	Accurate ground reaction force	Limited footstep coverage (one footstep per plate)
Pressure mat	C/NC	Capture ground reaction force distribution	Dense sensor array; Limited area of coverage
Cameras	C/NC	Interpretable body motion; Large coverage	Require clear line of sight; Privacy concerns
Wearables (IMU)	NC	Versatile movement capture; convenient	Require carrying devices; Limited body motion coverage
Pressure insole	NC	Capture ground reaction force distribution	Discomfort and inconvenience due to device wearing
Acoustics	NC	Contactless; Large coverage	Noisy signals; Privacy concerns due to voice recording
RF/Wi-Fi	NC	Contactless; Robust to obstructions	Noisy signals; Interference with the surroundings
Floor vibration	NC	Low cost, Contactless; Large coverage; Privacy friendly	Noisy signals; Sensitive to the floor and person contexts

*C/NC: Clinical/Non-Clinical setting

B. Prior Work on Floor Vibration-based Gait Health Monitoring

In the domain of gait health monitoring via footstep-induced floor vibrations, existing studies primarily serve exploratory purposes, with evaluations predominantly centered on feasibility assessments. Prior studies involve noise filtering and occupant detection, temporal gait parameter estimation, footstep localization, and left-right balance estimation. We provide an overview by categorizing these studies and summarize the main idea behind these methods.

One challenge when using footstep-induced floor vibration for occupant characterization is the mixture of environmental noises and human-induced vibration signals. To mitigate this issue, previous research has employed noise reduction techniques, such as low-pass filters and Wiener filters, combined with classification models to differentiate between footsteps and other signals. The approach involves filtering out noises and identifying peaks associated with footstep events in the time domain [25]. Then, we leverage the frequency domain of the vibration signals to classify human footsteps out of the other noises [24].

After reducing noises, existing studies have looked into temporal and spatial gait parameter estimation, such as the step time, step length, and cadence [20], [21], [33]–[35]. This is typically accomplished through the application of a peak-picking algorithm on the raw signal and/or the continuous wavelet transform (CWT) spectrum, followed by the selection of the range containing the predominant signal energy.

In addition, prior work also succeeded in estimating the cumulative ground reaction force and the left-right balance using the signal energy normalized by the person-to-sensor distances [36], [37]. To address the challenge that the vibration wave attenuates as the person walks away from the sensor, the method leverages the footstep localization algorithm to compensate for the energy loss due to attenuation. With the estimated gait forces per footstep, the study assessed the left and right foot balance by taking the ratio of the estimated forces.

Moreover, in order to overcome data misalignment challenges due to the variations of walking paths, floor types, shoe types, sensor locations, the number of walkers, and more, our prior work has explored models based on transfer learning to automatically align the data across various domains [21], [38], [39]. The main idea is to learn a mapping function that transforms the data from one domain to another in an unsupervised setting. This will make our model more adaptive across various real-world scenarios.

However, the existing studies have several limitations, including 1) lack of detection of gait abnormalities, and 2) overlook the challenge of person-to-person variations. In this paper, we bridge this research gap by tackling these limitations through the development of the first person-invariant gait abnormality detection system using footstep-induced floor vibrations. Our work pushes the boundaries by bridging the engineering developments on floor vibration sensing with real-world medical practices.

III. METHOD

In this section, we introduce the framework of gait abnormality detection through footstep-induced floor vibrations. As shown in Figure 1, our framework consists of three modules: 1) Gait Data Acquisition, 2) Gait Feature Extraction, and 3) Gait Abnormality Detection. In the following subsections, we present the purpose, components, and algorithms in each of the three modules, respectively.

A. Gait Data Acquisition

We develop a low-cost, maintainable floor vibration sensing system for gait data acquisition and pre-processing the footstep-induced floor vibration signals through noise filters and footstep detection algorithms.

The floor vibration sensing system leverages multiple geophone sensor nodes connected with a self-designed printed circuit board (PCB) for signal amplification and analog-to-digital conversion (ADC) as shown in Module 1 of Figure 1. Each independent sensor node is individually robust and simple by design based on our prior work [40]. Geophone sensors have many advantages in human gait monitoring. The sensing system can be operated by low power (compatible with a 2200mAh 11.1V lipo battery) and is low cost. They are designed to restart quickly and without intervention after errors such as temporary power loss, signal degradation, overflowing buffers, and minor node damage. Compared to accelerometers and microphones, a geophone's frequency range aligns well

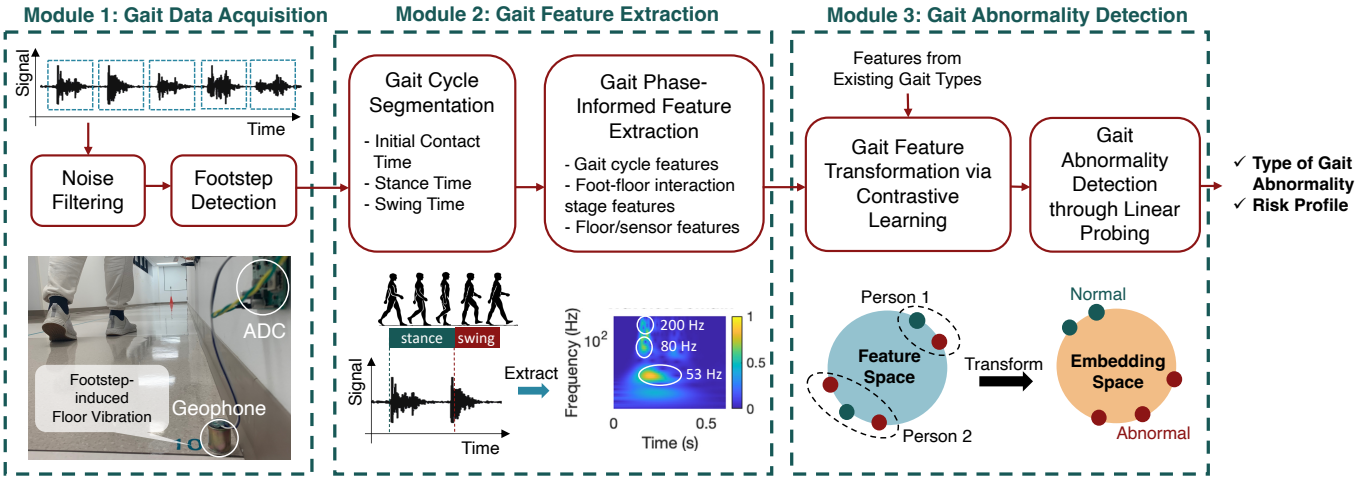

Fig. 1. Overview of our floor vibration-based gait abnormality detection framework, which consists of three modules: 1) gait data acquisition, 2) gait feature extraction, and 3) gait abnormality detection.

with the effective range of human gait-induced floor vibrations, and it is perceived as more privacy-friendly for in-home usage than microphones. The sensing range of each node depends on the floor properties. When tested on a wooden floor with 500× amplification, each sensor covers up to a 20-meter radius. All participants have provided written informed consent and/or assent and all experiments are conducted in accordance with the approved protocols under relevant guidelines and regulations of Stanford Research Compliance Office (RCO) with approved protocol number IRB-55372 on 08/20/20 by the Stanford Institutional Review Board (IRB).

The noise filtering sub-module aims to reduce noises that are not relevant to human gait. We use a lowpass and a Wiener filter. The lowpass filter serves the purpose of eliminating high-frequency noises due to electrical interference and/or mechanical vibrations. The lowpass filter is configured with a threshold of 500 Hz. This setting allows us to retain the vital gait information present in floor vibration data, specifically the range from 5 to 250 Hz, as demonstrated in prior work [19]. In addition, the Wiener filter serves the purpose of reducing background environmental noises [41]. This filter takes in a 3-second signal segment containing only background noise, leverages its frequency spectrum, and filters out the noise spectrum from signals.

The footstep detection sub-module aims to detect individual footsteps from the continuous vibration signals collected by our sensing system. Our algorithm is developed based on peak identification within the wavelet coefficients. We apply wavelet transformation to the entire signal using the Morlet wavelet, a widely adopted wavelet known for its computational efficiency and suitability for capturing time-varying, non-stationary signals [42]. Given that footstep-induced vibration signals possess an inherently impulsive nature due to the short foot-floor contact duration, we narrow our focus to the natural frequency range typically found in floor structures (5-30 Hz) when analyzing the wavelet coefficients. Within this range, we identify the peaks in the wavelet coefficients as footstep impulses. In addition, since footsteps tend to occur in repetitive patterns in the vibration signals as a person continues to walk, we set the minimum number of consecutive impulses to three. This enables us to distinguish footsteps from other human-induced impulse signals, such as items dropping and doors opening/closing, which has been evaluated in our prior work with satisfactory classification accuracy [24].

B. Gait Feature Extraction

We extract gait features from floor vibration signals by first segmenting the gait cycle into three foot-floor interaction stages and then extracting features from these stages based on the physical relationship between footstep force and floor vibrations.

1) Gait Cycle Segmentation: In medical practices, human walking is assessed based on the segmentation of gait cycles. A gait cycle is a fundamental unit of human walking, defined as the time interval of locomotion between two successive strikes (initial contacts) of the same foot [43]. A typical gait cycle consists of two phases: 1) the stance phase when the foot is in contact with the floor and 2) the swing phase, where the foot lifts off from the floor and swings in the air. In this study, we focus on three foot-floor interaction stages of a gait cycle based on the occurrence of various types of gait abnormalities, including initial contact, stance phase, and swing phase.

In this sub-module, we extract the time between initial contact and foot-off, which divides a gait cycle into segments that correspond to stance and swing phases as shown in Module 2 of Figure 1. To begin with, we identify gait cycles by grouping the previously detected footsteps detailed in Section III-A. Since a standard gait cycle consists of one foot's strike and the subsequent strike of the opposite foot, we pair each consecutive left and right footstep to form a gait cycle group. Subsequently, we determine the timing of foot strikes and foot-offs to segment the gait cycle. To achieve this, we compute the sum of wavelet coefficients within the dominant frequency ranges of the floor (around 23 Hz for our testing floor) and the higher frequency range in which the foot strike-induced vibration signals typically lie. This calculation

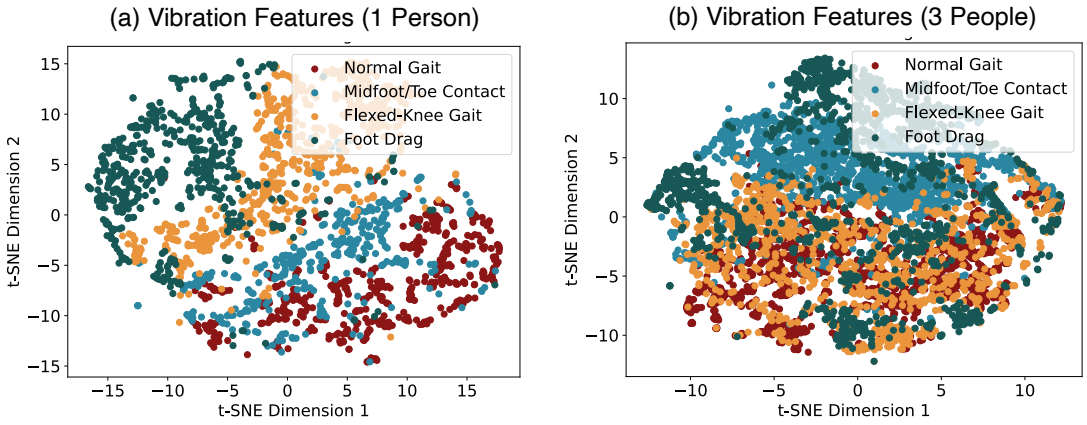

Fig. 2. Visualization of extracted features through 2D t-SNE feature plots. (a) shows that the vibration features from 1 person show a clear distinction between different gait types, indicating the effectiveness of our features. (b) shows the feature from 3 people, which has significant overlapping between people, signifying the challenge posed by the individual differences.

yields two distinct time series: one for detecting foot strikes, associated with higher frequencies, and another for foot-offs, related to lower frequencies, following the characterization from our prior work [21]. We then leverage peak detection algorithms on these resulting wavelet coefficient time series to identify specific times of interest.

To extract the initial contact time, when the higher frequencies first become prominent, we apply a reverse sliding window technique, starting from a peak and proceeding to a valley. This approach enables us to pinpoint the time when the vibration signal amplitude begins to rise, signifying the moment of foot strike. In addition, the foot-off time is detected as the peak of the lower-frequency component because it is the time when the damped free vibration starts to attenuate the signal. Finally, each gait cycle is divided into segments based on these foot strike and foot-off times. This division allows us to identify the various foot-floor contact stages accurately.

2) Gait Phase-Informed Feature Extraction: With the segmented gait cycle, we extract features from three perspectives: 1) gait cycle features, 2) foot-floor interaction stage features, and 3) floor/sensor features. In particular, the foot-floor interaction stage features are tailored to the various types of gait abnormalities at initial contact, during the stance phase and swing phase, respectively. The floor/sensor features provide the context of the sensor configurations on the floor to enhance model generalizability across various floors and sensor layouts. The features we extracted are summarized below.

- **Gait Cycle Features.** The gait cycle features aim to describe the general characteristics of a data sample. These features include the left/right foot indicator, signal energy, and variance every 0.1 sec, as well as the mean power spectral density every 10 Hz of each gait cycle. These features provide an implicit description of the gait cycle that potentially informs the type of gait abnormalities.
- **Initial Contact Features.** The initial contact features include the mean wavelet coefficients of each 10 Hz frequency range for each 0.2-sec range around the initial contact time. The frequency resolution is chosen based on the observation of the minimal interval between various modal frequencies. For the temporal range, we focus on the wavelet coefficients 0.1 sec before and after the initial contact time to allow imprecise estimations from the previous segmentation algorithm while minimizing the feature dimension for efficient model training.
- **Stance Phase Features.** The stance phase features include the proportion of lower frequency components ($\leq$ 60 Hz) among the overall frequency spectrum during the segmented stance phase. The threshold is chosen based on preliminary testing on the floor structure, where we observe that the 1^{st} and 2^{nd} modes of the floor are both below 60 Hz. To mitigate error propagation from the gait cycle segmentation, we incorporate a 0.1-sec margin before and after the stance phase when computing the frequency spectrum.
- **Swing Phase Features.** The swing phase features include the number and amplitude of the peaks within the high-frequency range (100-200 Hz) in wavelet coefficients within the swing phase. The frequency range is chosen based on the observed typical frequencies induced by foot drags. Similar to the stance phase features, we incorporate a 0.1-sec margin before and after the swing phase to allow up to 50% estimation error from the previous step.
- **Floor/Sensor Features.** To provide the context of the floor and sensing system, we include the sensor number and gait cycle number as one-hot-encoded features. The sensor number refers to the label of each sensor, which helps to distinguish the signal variations due to common hardware flaws during manufacturing and the coupling effect between individual sensors and the floor. The gait cycle number informs the location of the footstep relative to the sensors, which helps to incorporate the heterogeneity in the floor properties and the variations of footstep-to-sensor distances.

These features are first concatenated and then processed through missing-data handling and standardization algorithms to compile a clean and complete dataset. The standardization algorithm normalizes the features to the same scale to minimize the biased penalty in gradient descent during training.

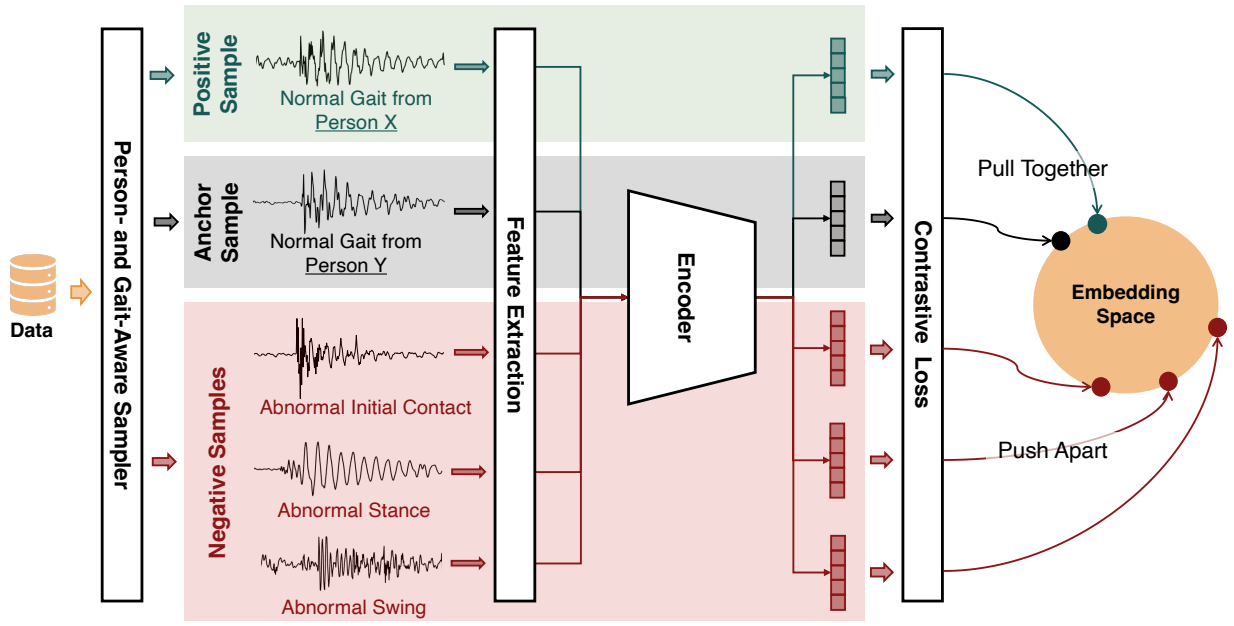

Fig. 3. An example of the contrastive feature transformation process through sampling and contrastive learning. For each anchor sample, we select a positive sample and multiple negative samples to train a feature transformation encoder through contrastive loss, which pulls together the anchor and positive sample (sample with the same gait type but from a different person) while pushing apart the anchor and the negative samples (samples with different gait types) in the embedded feature space.

After that, we apply a feature selection model to filter out the less effective features and preserve the ones that provide the most gait-related information. Specifically, we use the random forest model to compute the feature importance using the training data and labels, which indicates the contribution of each feature to the model's predictive performance. We narrow down the feature number to 64 to enable efficient training, the pattern of which concerning each gait type is shown in Figure 2.

While the extracted features represent gait abnormalities, they also contain information on between-person variations during walking, leading to difficulty generalizing to new patients. As we observe from Figure 2a, the extracted vibration features from 1 person are effective for gait abnormality detection as there are clear distinctions among various types of gait. However, when visualizing the features from 3 people (see Figure 2b), we find that the features from various gait types are less separable due to the difference among various people's gait. For example, there are significant overlaps between normal gait (red dots) and flexed-knee gait (yellow dots) because each person has a unique gait, which is reflected through the floor vibration data as discussed in our prior work on floor vibration-based person identification [30], [31]. As a result, a person's normal walking gait may share similar features as another person's flexed-knee gait, making it challenging to detect gait abnormality for people who have no pre-recorded data. This calls for a person-invariant gait abnormality detection method that can be generalized to newly observed people, which is introduced in the following Section.

C. Gait Abnormality Detection

With the extracted features, we develop a person-invariant gait abnormality detection method by first transforming the features into an embedding space such that the difference among various gait types (e.g., normal vs. abnormal gaits) is preserved while the difference among various individuals is reduced. After the transformation, we leverage a linear probing to compare the gait abnormality templates with the newly observed data to detect abnormal gait.

1) Gait Feature Transformation via Contrastive Learning: To achieve the person-invariant goal, gait feature transformation involves two steps: the development of a person- and gait-aware sampler and a contrastive learning model [44]–[46]. An example of feature transformation through sampling and contrastive learning is described in Figure 3.

During person- and gait-aware sampling, each gait cycle sample is defined as an anchor. For each anchor, the sampler selects a corresponding positive sample (which belongs to the same gait type as the anchor but from a different person) and multiple negative samples (which belong to different gait types from the anchor) from the training data. The positive sample is selected from a different person from the anchor to force the model to recognize the difference between individuals and close this gap. The negative samples are selected from each of the different types of gait abnormality regardless of the person, enabling the model to distinguish between different gait abnormality types.

The contrastive loss we use is based on InfoNCE [46], where NCE stands for Noise-Contrastive Estimation, which is based on the principle of maximizing mutual information between positive pairs and minimizing it between negative pairs. InfoNCE has shown promise in time series data and has superior performance when working with siamese encoders (i.e., same encoder for different gait samples) for similarity comparison [47]–[49], which aligns with the goal of our abnormal gait detection problem.

The encoder described in Figure 3 is a 3-layer neural network, with a hidden dimension of 256, which is a common choice for many basic machine learning tasks. It can capture non-linear representations in the data while being efficient and simple enough to mitigate over-fitting. The depth and the size of the neural network are chosen based on a grid search among various sets of model parameters.

2) Gait Abnormality Detection through Linear Probing: We now aim to answer the question: *Are the representations learned by contrastive loss useful for the downstream classification task?* To study the quality of the learned embeddings, we follow the widely used linear probing protocol [46], [50]–[53], where we train a simple logistic regression model on top of the frozen encoder network and use test accuracy as a proxy for representation quality. This linear probing protocol is commonly used to evaluate contrastive learning methods. This is because, if the learned representations allow a simple linear classifier to achieve good performance on a specific task, then it suggests that the representation has captured general and discriminative features about the data.

In this study, linear probing is also used as a downstream classifier for gait abnormality detection. To achieve this, we first train three separate linear classifiers using the learned embeddings from each of the three foot-floor contact stages, and then test the classifiers on the embeddings from the newly observed samples. We integrate the embedding from multiple sensors using a weighted loss during linear probing for gait abnormality detection [54]. The weight is assigned based on the signal-to-noise ratio of the data - closer sensors typically have higher weights as their signal amplitudes relative to the noise amplitudes are significantly higher than the sensors that are farther away. During testing, the embeddings from the newly observed samples are computed by feeding through the trained encoder, which transforms the new features into the embedding space. Such siamese architecture (i.e., perfectly symmetrical train and test encoders) has been shown to work well with contrastive loss in numerous existing studies [55], [56].

The detection results are whether or not a person has gait abnormality during initial contact, stance phase, and swing phase, respectively. To understand the specific type of gait abnormalities, we focus on four typical gait types in this study, including normal walking, gait with flatfoot/toe contact, flexed-knee gait, and gait with foot drag, and predict the type to which the test sample is most likely to belong. All the results are produced for individual gait cycles to preserve the distribution of a person's gait and are further used for visualization of personalized gait abnormality risk profiles, which is described in Section IV-C.

IV. Evaluation and Results

To evaluate our approach, we conducted a real-world walking experiment with 21 human subjects and collected 62,404 gait cycles. In this section, we first introduce the evaluation setup, and then discuss the overall results of abnormal gait detection, and finally create diagrams to visualize the personalized gait abnormality risk profiles.

A. Evaluation Setup

The evaluation setup consists of the experiment setup and the data preparation. We first introduce the field experiment setup (including equipment, procedures, and data flow), and then describe the data preparation process (including labeling, train-test splits, and model parameters).

The experiment involves two sets of sensors: 1) four geophones mounted on the surface of the floor for vibration data collection and 2) a Vicon Motion Capture (MoCap) system to capture the walking posture of each subject for ground truth. The vibration sensors are placed at the side of the walking path as shown in Figure 4a,c, with a sampling frequency of 500 Hz. The analog signals are converted into digital signals and amplified to a range of -5 to 5 V through Vicon Locker Lab to enable synchronization between camera and vibration data [57]. The data are stored on a desktop server to mimic the actual data flow in home environments. The floor vibration system can also be adapted to multi-room settings using the wireless sensor network with Wi-Fi-based data transmission developed in our prior work [58].

During the experiment, each of the 21 participants walked across a 7-meter long walkway for 30-40 trials according to four types of gait to simulate the normal and abnormal contacts, including normal walking, abnormal initial contact (flatfoot or toe contact), flexed-knee gait, and foot dragging (see Figure 4b). Each trial typically consists of 6-8 gait cycles. The participant characteristics are summarized in Table II. All participants have provided written informed consent and/or assent and all experiments are conducted in accordance with the approved IRB (IRB-55372) by the Stanford Institutional Review Board as mentioned in Section III-A. During each walking trial, the participant first stands still and then accelerates to reach a constant speed. The walking speed is chosen by the walker during normal walking and qualitatively instructed during abnormal walking by medical experts. The specific instructions provided by the medical experts include 1) qualitative characteristics of each gait abnormality type (e.g., toe strike instead of heel strike), 2) muscle activation sequence and status (e.g., relaxing the calf muscles while contracting the quadriceps during leg swings), and 3) thresholds for the desired ranges of motion in hip, knee, and ankle joints. These instructions lead to more realistic abnormal gait while keeping the natural variability of different people within each gait abnormality type. When reaching the end of the walkway, the subject decelerates and stops walking. The operation of the MoCap system involves 16 markers attached to the subject's lower limbs in accordance with the Plug-in gait lower body model in Vicon MoCap System [57], producing records of locomotion during walking with a sampling frequency of 100 Hz. The data are manually labeled by a trained expert. The labels include the "initial contact" and "toe off" time for each gait cycle (see Figure 4d for a sample gait cycle in vibration with decomposed frequency components). Among all the collected data, samples with insufficient data quality or unrealistic abnormal walking patterns are discarded as determined by the medical experts. Overall, there is a total of 62,404 labeled gait cycles summarized in our dataset, each

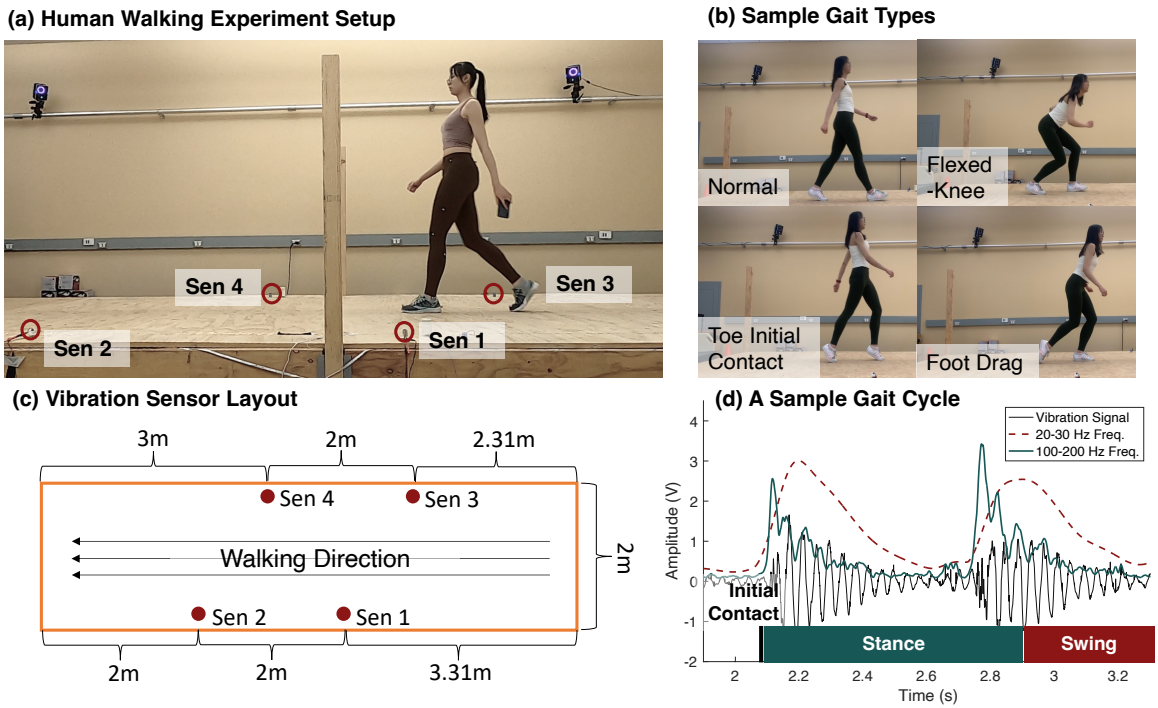

Fig. 4. Real-world walking experiment setup. (a) shows the human walking experiment in action with the vibration sensors mounted at the edge of the walkway. (b) presents examples of gait types during the experiment. (c) describes the vibration sensor layout and the dimensions of the walkway. (d) shows a sample gait cycle with initial contact, stance phase, and swing phase.

further attached with a person label and a gait type label.

For data preparation, we split the train and test data through random selections among subjects. In our evaluation, ∼80% of the subjects belong to the training set (rounded to 17 people) and the rest belong to the test set (rounded to 4 people). While the majority of the existing studies using floor vibration for human monitoring use footstep-level train-test split [19], [30], [38], our evaluation uses the subject-level split, which has multiple benefits and better reflects the real-world scenario. The subject-level split allows evaluation of our model performance for previously unobserved people, which shows the robustness and generalizability of the model to new subjects. In our problem, it validates the effectiveness of the person-invariant feature transformation. In addition, the subject-level split also aligns with the practical setting where the need for gait abnormality detection typically comes from new patients. The reported performance of the models is computed by averaging 10 random selections to represent the overall performance, assuming that the variance of 10 random tests is representative of all possible train-test combinations.

To demonstrate the effectiveness of our contrastive learning approach, we use a baseline model (a 3-layer feedforward neural network encoder) that replaces the contrastive loss with cross-entropy loss in Figure 3. The model parameters are tuned through a grid search algorithm via 5-fold cross-validation on the training data, and the selected parameters are used for model testing. The tuned parameters include the learning rate, the number of neurons in the hidden layers, the dimension of the embedding space, the number of negative samples for each anchor, and the batch size. Both the baseline model and our model are optimized through parameter tuning to produce a fair comparison.

B. Results on Abnormal Gait Detection

Overall, our method achieves an 85%, 90%, and 95% mean test accuracy in detecting gait abnormalities for each gait cycle during initial contact, stance phase, and swing phase, respectively (see Table III). Compared to the baseline models without person-invariant contrastive learning, our method has a $1.3\times$ to $2.1\times$ error reduction in detecting abnormalities among various foot-floor interaction stages. The deep learning models (LSTM and CNN) have similar or worse performance than many of the statistical learning models, are more unstable during training, and require significant effort on hyperparameter tuning due to the higher model complexity.

It is worth noting that all evaluation results are described for individual gait cycles instead of a walking trial or a person. This provides better granularity of information and allows further integration of predicted outcomes as a person typically has multiple walking trials and each trial consists of multiple gait cycles. For example, our prior work on gait-based disease progression tracking aggregates the information from individual gait cycles for each walking trial and then for a person's overall gait patterns through hierarchical models, which has shown to improve the prediction accuracy from ∼70% to ∼90% [59]. In addition, predicting individual gait cycles enables flexibility in downstream tasks such as estimating the distribution of normal and abnormal footsteps within a trial or a person.

1) Effectiveness of the Person-Invariant Feature Transformation: We evaluate the effectiveness of the person-invariant

TABLE II
PARTICIPANT CHARACTERISTICS (N=21). PARAMETERS ARE GIVEN AS MEAN ± STANDARD DEVIATION OR COUNTS IN CASE OF SEX.

Age (yrs)	Gender (F/M)	Height (cm)	Weight (kg)	Left Leg Length (cm)	Right Leg Length (cm)	Knee Width (cm)	Ankle Width (cm)
24.6 ± 3.0	11 / 9	169.4 ± 9.9	63.8 ± 14.2	87.7 ± 6.6	87.6 ± 6.4	9.8 ± 1.2	7.1 ± 0.5

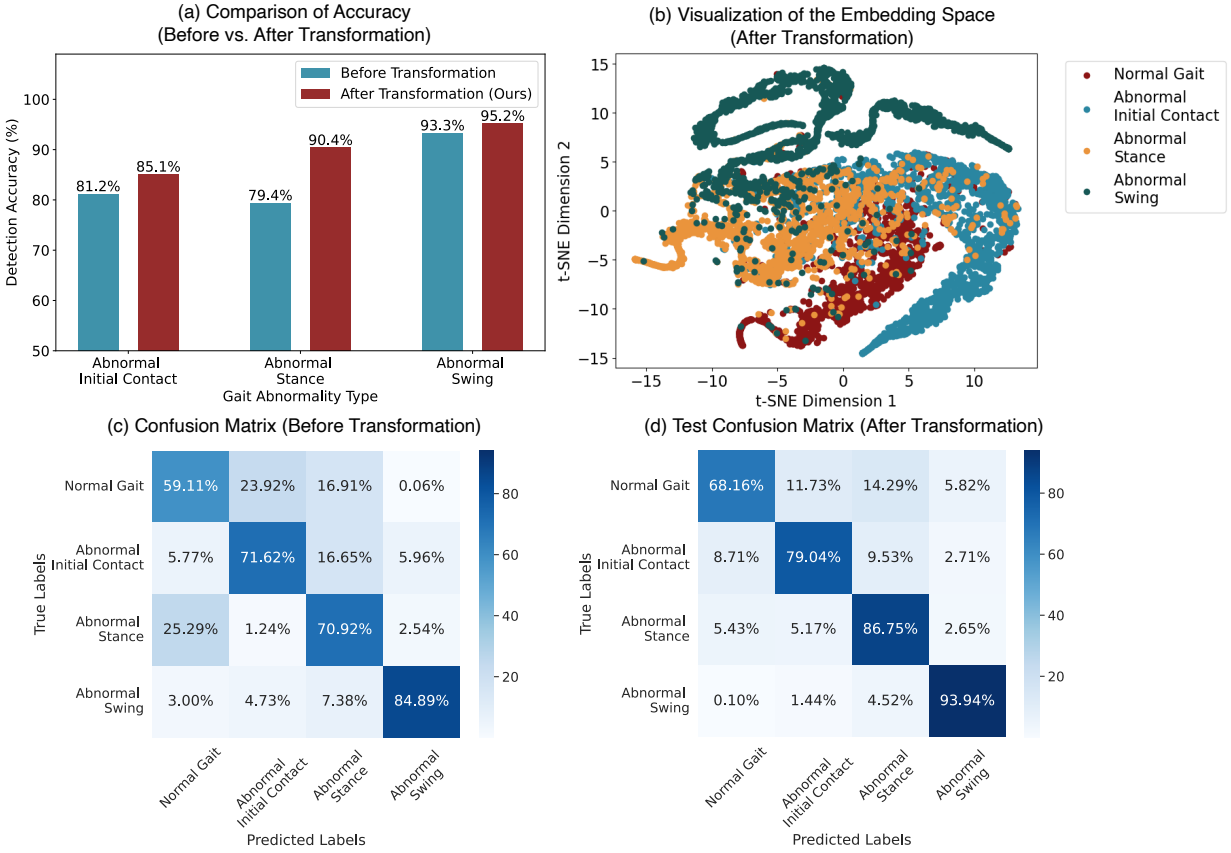

Fig. 5. Effectiveness of the person-invariant feature transformation. (a) Our method outperforms the baseline model in gait abnormality detection for abnormalities during initial contact, stance phase, and swing phase, respectively. (b) 2D t-SNE plot of the embedding space (from all test participants) shows less overlapping among different gait types compared with Figure 2b, demonstrating the effectiveness of the feature transformation. (c)(d) Our method has a 20.9% improvement in F1 score over the baseline when comparing the test confusion matrices for gait type classification.

TABLE III
FLOOR VIBRATION-BASED GAIT ABNORMALITY DETECTION ACCURACY

Model	Abnormal Initial Contact	Abnormal Stance	Abnormal Swing
Naive Bayes	0.6868	0.6899	0.9132
Quadratic Discriminant Analysis	0.7079	0.7384	0.9316
K-Nearest Neighbors	0.8049	0.7964	0.9301
Decision Tree	0.7535	0.7709	0.9177
Random Forest	0.7823	0.7049	0.9269
AdaBoost	0.8006	0.7848	0.9280
Multi-Layer Perceptron	0.8120	0.7941	0.9328
Long-Short-Term Memory (LSTM)	0.8171	0.8142	0.7025
Convolutional Neural Network (1D-CNN)	0.7592	0.6570	0.6513
Our Method	**0.8510**	**0.9041**	**0.9521**

contrastive learning through an ablation study to compare the same model framework before and after the feature transformation, the results of which are summarized in Figure 5. After the feature transformation, Our method has improved the performance for all gait abnormality types. Figure 5b visualizes the t-SNE plot of the embedding space from 17 randomly selected training subjects. The plot is generated by compressing the high-dimensional embeddings into 2D space by preserving their relative location in the space. In this plot, we observe a clear separation of the feature embeddings among various gait types (including normal and abnormal gait) after the feature transformation. Compared to the similar plot (Figure 2b) presented in Section III-B, there is significantly less overlapping of features in the embedding space than the feature space. This establishes the efficacy of our person-aware sampling and contrastive learning approach for person-invariant gait abnormality detection. The detailed discussion of the individual difference is presented in Section V-A.

In addition to abnormality detection, we also evaluate the performance of our method in classifying various gait types, including normal walking, flatfoot/toe contact, flexed-knee gait, and gait with foot drag. Our method has an average of 0.81 test F1 score per gait cycle with a 20.9% improvement

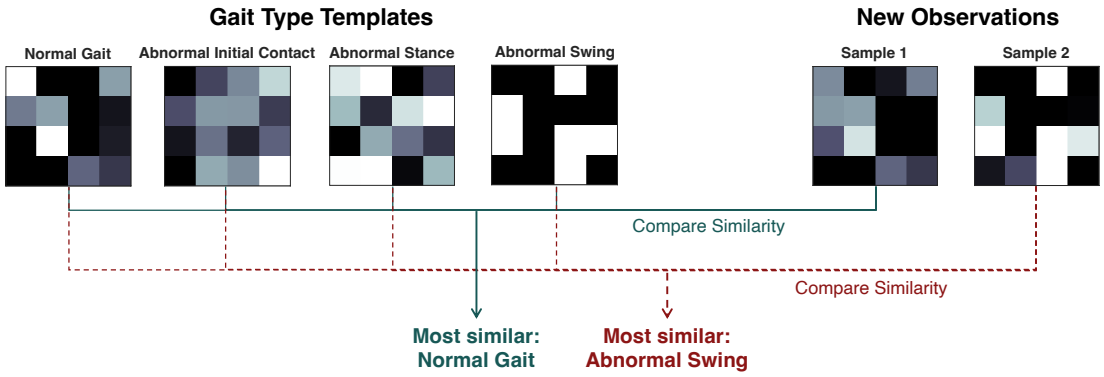

Fig. 6. Comparison between the gait type templates and the new observations in the embedding space. Each gait type has a unique pattern in the transformed embedding space. The new observations are first transformed to this embedding space and then compared with the templates through similarity measures to identify the most similar gait type.

over the model before transformation. Specifically, our method has better precision in all gait types (seen when comparing the percentages along the diagonal line between Figure 5c and d). Similar to the abnormality detection results, both baseline and our method have the highest precision for classifying abnormal swing as it is most obviously distinct from normal walking. The lowest precision lies in the normal gait prediction due to potential variations among gait cycles during walking. The false positive rate (the first row of Figure 5d) in detecting abnormalities is higher than the false negative rate (the first column of Figure 5d). This can be caused by the accelerating and decelerating gait cycles during normal walking. For example, when a person accelerates, the center of mass tends to lean forward and the footstep force has fewer horizontal components than normal heel strikes and tends to be more similar to toe strikes. The evaluation results establish the efficacy of our method for gait abnormality detection and gait type classification.

C. Results on Personalized Gait Abnormality Risks

We define and develop personalized risk profiles to visualize each person's gait abnormality risks based on the distribution of similarity measures between the gait type template and the embeddings of the new observations.

1) Templates of Normal and Abnormal Gaits: To visualize the effectiveness of the trained embeddings for person-invariant gait type classification, we develop templates for each gait type and obtain the similarity measures between templates and test samples. The gait type templates are computed by averaging all the training embeddings of the same gait type among multiple people to determine the centroids of each gait type. Then, we compare the similarity between the embeddings from new observations and the embedding templates from the training foot-floor contact data through a similarity metric to predict the type of gait abnormalities. For similarity comparison, we choose cosine similarity as the metric, because it is scale-invariant and captures the directional difference between vectors in an embedding space. In addition, cosine similarity is more interpretable as it produces values between -1 and 1, where 1 indicates identical directions, 0 indicates orthogonal (unrelated) directions, and -1 indicates exactly opposite directions. These values are intuitive and can be easily interpreted. Figure 6 shows an example of the templates and test samples in the embedding space.

In this example, each gait type template has a unique pattern in the transformed embedding space, representing the significant difference between normal and abnormal gait, as well as the difference among various abnormality types during initial contact, stance phase, and swing phase. When computing the similarity between new observations and the templates, sample 1 has the highest similarity with the normal gait, and sample 2 is more similar to the gait with abnormal swing (i.e., foot drag), which can be observed in the patterns shown in Figure 6.

2) Personalized Profile of Gait Abnormalities: We develop personalized profiles of gait abnormalities based on the cosine similarity between a person's gait samples and the gait type templates. We apply a scaling function to convert the similarity to risk values from 0 to 1. During this process, we quantify the risk through the scaled metric that is representative of the relative risk of each type of gait abnormality. We assume that the risk values for each gait are independent and thus do not add up to a fixed sum. This is because the types of gait are not mutually exclusive, meaning that a person may have multiple types of gait abnormalities at the same time.

Figure 7 shows the personalized profiles of two test subjects. Overall, we observe variations in both people's gait, which are indicated by the scattered dots along the horizontal axis. In this example, the distribution of gait cycle similarities of person 1 has a smaller variance than that of person 2 (shown as the shorter length of the box plot for person 1), meaning that person 1 has a more consistent gait and fewer variations. When comparing each abnormality type, person 1 has a lower risk than person 2 in general, as observed in the lower mean value in the box plot. In particular, person 2 has a higher risk of abnormal initial contact as the similarity with the abnormal initial contact template is high and the similarity with the normal gait template is lower and varies significantly. These profiles provide insights into a person's overall gait pattern as well as the distributions among gait cycles. Compared with the existing clinical reports that focus on a specific gait cycle,

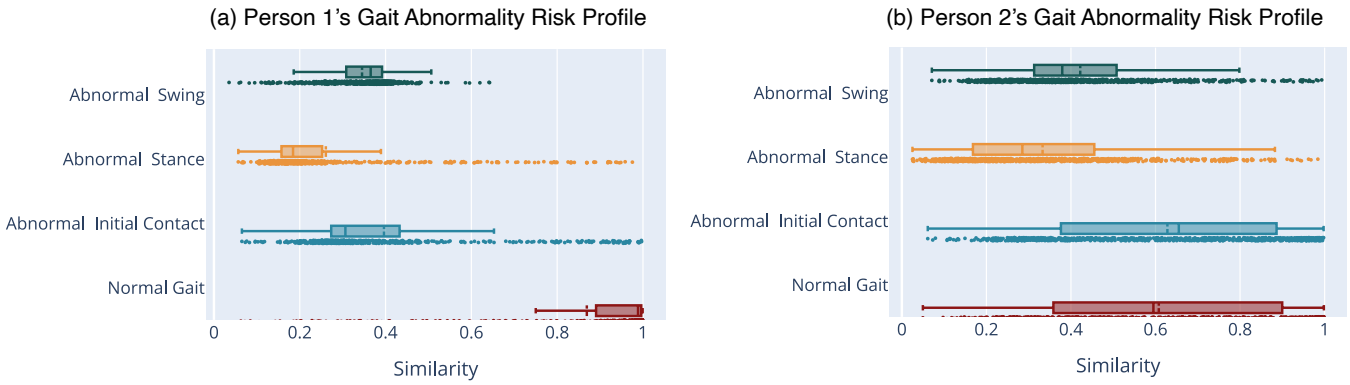

Fig. 7. Personalized gait abnormality risk profile for two test subjects with their natural gait: (a) shows person 1's profile and (b) shows person 2's profile. Each dot represents a gait cycle. Person 2 has higher variability among gait cycles than person 1 and has significantly high risks of abnormal initial contact based on the similarity to the gait type templates.

our method provides information on the variation over many gait cycles, which can reduce the sampling bias and bring in new opportunities for long-term monitoring in daily life.

V. Discussion

We discuss the effect of individual differences on gait abnormality detection and show that our feature transformation can effectively be generalized for various people, gait types, and sensor locations.

A. Discussion on Person-to-Person Variations

To address the challenge posed by individual differences in gait patterns, our method transforms the features into an embedding space where individual differences are reduced and gait type differences are amplified. We visualize the features/embedding before and after transformation with respect to various people to discuss the effect of person-to-person variations, as shown in Figure 8. In the feature space (upper row of Figure 8), there is a significant difference among different people's data within the same type of gait (i.e., clear separation in different colors for each gait type). In contrast, the embedding space plots (lower row of Figure 8) show that there is little separation between different people under the same type of gait, meaning that the transformation is effective in reducing between-people differences.

Based on the gait type templates, we compare the similarity of the feature embeddings across gait types and people and visualize them in Figure 9. When plotting the feature embedding by gait type (see Figure 9a), we observe that the embeddings from the same type of gait have significantly higher similarity than the embeddings across different gait types, forming a bright diagonal line in the matrix (brighter color means higher similarity). It is worth noting that there are also many similar samples between the normal gait and the gait with flatfoot/toe contacts, which means that it is more difficult to distinguish between these two gait types than between the others. This may be because some individuals' normal gait may be more similar to flatfoot contact and there are significant variations among various gait cycles as a person accelerates or decelerates. On the other hand, there are several samples in foot drag that have higher similarity with normal gait. This is because the foot drag may not be present in every gait cycle as a person walks even when the person has gait abnormalities during the swing phase. To this end, we combine the information from individual gait cycles to a representative gait pattern of a person, which is visualized through personalized gait abnormality risk profiles (which will be introduced in Section IV-C). In Figure 9b, we observe that there are evenly distributed bright areas when plotting the similarity by people. This means that the same person's gait cycles do not always have similar features except for the samples within the same gait type. The brighter areas of the diagonal line also indicate that samples from the same gait type of multiple people become similar after the feature transformation, meaning that the individual differences have been reduced.

In addition to the individual difference in walking patterns, the number of training samples from each individual may also differ, which can lead to over-fitting due to the imbalanced data distribution. This problem can be addressed by re-sampling the data using techniques such as randomization, Monte Carlo, or Bootstrap in future work.

In summary, the results above demonstrate that our method is robust to person-to-person variations, resulting from the person-invariant feature transformation. Our method has successfully reduced the individual differences and distinguished among different gait types (including normal gait and various types of abnormal gait).

B. Discussion on Gait Type Variations

The detection of abnormal initial contact, abnormal stance, and abnormal swing has an F-1 score of 0.8513, 0.9174, and 0.9494, respectively, as shown in Figure 10. This means the performance of our method varies with the types of gait abnormalities. Among all the gait abnormality types, we observe that the abnormal swing has the highest accuracy for both baseline models and our method. This is because there are obvious differences in vibration signals induced by additional foot-floor contacts during the swing phase when compared to no contact in normal walking. Also, the false negative rate is low for abnormal swing detection, which

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2024.3413815

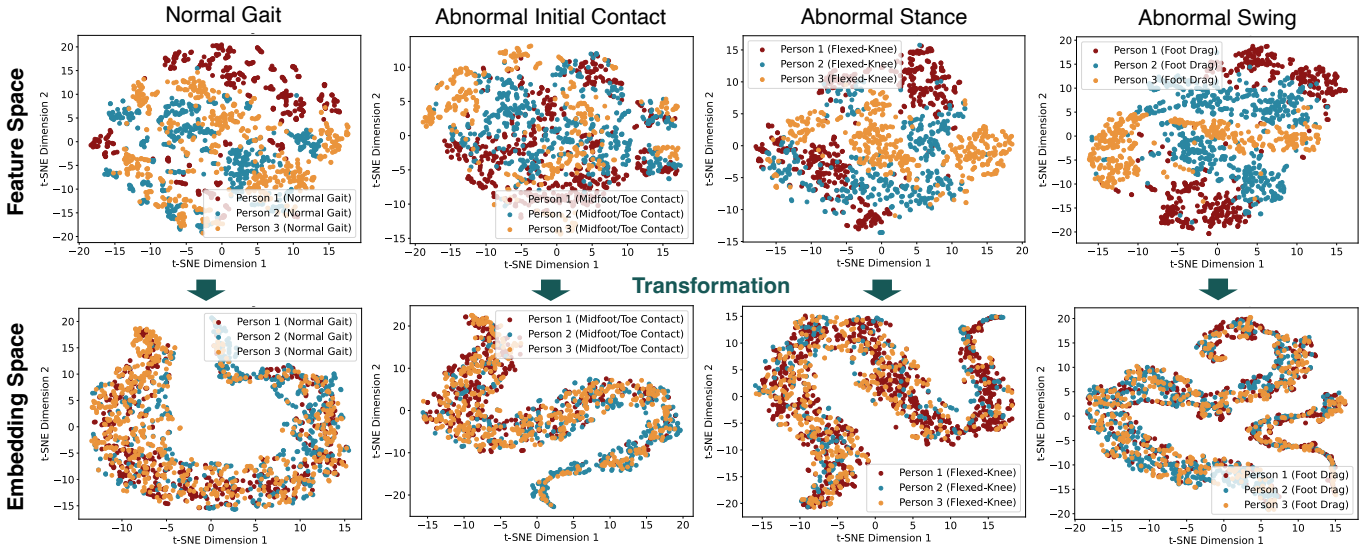

Fig. 8. Comparison of t-SNE feature plots before and after the feature transformation from 3 randomly selected test subjects (dots with different colors). For each gait type, there is a significant difference among various people's gait before the transformation (upper row, representing the feature space); After the transformation, such difference is significantly reduced as people's embeddings align with each other (lower row, representing the embedding space).

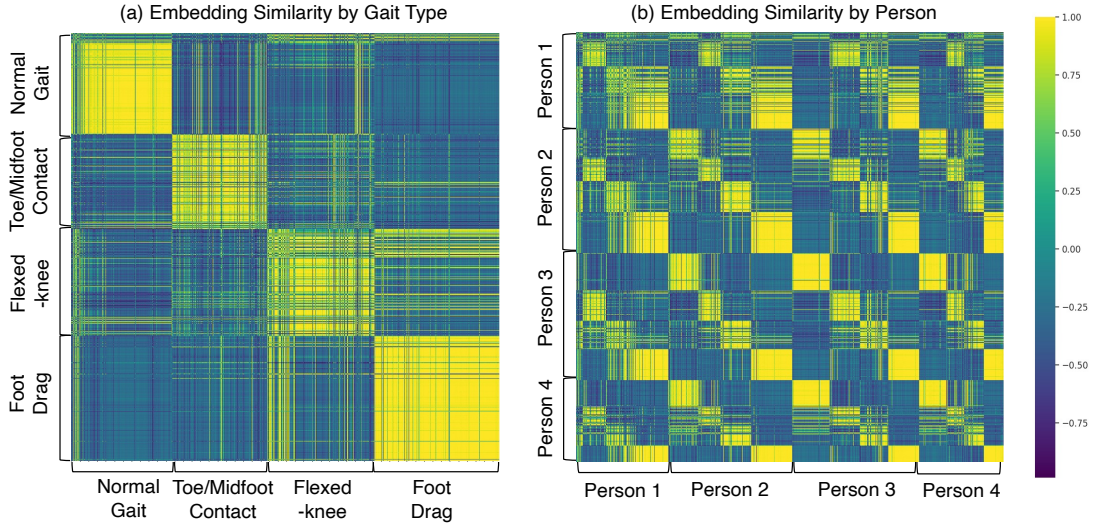

Fig. 9. Cosine similarity of the transformed features visualized by (a) gait type and (b) person, with a scale from -1 to 1 (including 8067 gait cycles in the test set). (a) shows that the transformed embeddings within the same gait type (along the diagonal) have higher similarity than the others. (b) shows that the transformed embeddings are invariant to different people as the high-similarity samples are evenly distributed among test individuals.

means the model is highly sensitive, and thus fewer cases of gait abnormalities are missed. Conversely, the abnormal initial contacts (i.e., flatfoot/toe contact) detection has the lowest accuracy for both models, mainly because of the short duration of initial contact (typically within 0.2 sec). We also observe that the number of false positives for abnormal initial contact is larger than that of the false negatives, which can be explained by the variations in angle and force combinations during normal walking. For example, some individuals tend to lean forward, which makes them more likely to have a flatfoot initial contact. By comparison, the abnormal stance detection has a higher false negative rate, which means some abnormal samples are regarded as normal. This may be because the weight translation speed of flexed knee gait can be similar to normal gait at the beginning of walking and is likely to decrease significantly after a certain distance due to muscle fatigue.

It is worth noting that our method only detects gait abnormalities within the scope of existing abnormal gait data. Since contrastive loss focuses on distinguishing the difference between normal and abnormal gait based on the training data, it may not generalize well to previously unobserved gait types. However, our method can be easily adapted to new abnormality types by adjusting the person- and gait-aware sampler to generate more negative samples (i.e., abnormal gait types) in Figure 3.

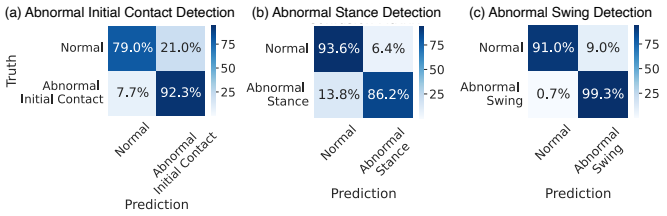

Fig. 10. The performance of gait abnormality detection across various gait types, including (a) abnormal initial contact, (b) abnormal stance, and (c) abnormal swing. Abnormal swing detection has the best performance among all types.

TABLE IV
EFFECT OF DIFFERENT SENSORS ON DETECTION ACCURACY

Sensor	Abnormal Initial Contact	Abnormal Stance	Abnormal Swing
Sen 1	0.8203	0.8591	0.9291
Sen 2	0.8338	0.8350	0.9348
Sen 3	0.8143	0.8412	0.9362
Sen 4	0.8426	0.8726	0.9425
All Sensors	**0.8510**	**0.9041**	**0.9521**

The analysis above shows that the performance difference among gait types can be explained by the unique characteristics of each gait type. In the future, we will explore other types of variations such as the walking speed, walking path, floor type, shoe type, and so on to transfer the person-invariant learning model into speed-, path-, floor-, and shoe-invariant models. Moreover, we plan to further balance the trade-off between false positives and false negatives and explore more types of gait.

C. Discussion on Sensor Location Variations

We evaluate the effect of sensor location on gait abnormality detection results, which are summarized in Table IV. Sen 1, Sen 2, Sen 3, and Sen 4 represent individual sensors labeled in Figure 4, respectively. From Table IV, we observe that Sen 4 has the best performance among all. This is because Sen 4 has the best signal-to-noise ratio (SNR) as it is located near the center of the walkway (i.e., closer to the majority of the footsteps) and right above the mid-span between two joists (i.e., more significant floor vibration there). While Sen 1 is also near the center of the walkway, it is placed on top of a joist where the vibration is less significant due to the higher stiffness. Sen 2 and Sen 3 have similar performances as their locations are symmetrical. Since they are placed closer to the two ends of the walkway, the SNR for footsteps landing at the other end is significantly lower than those of Sen 4. This analysis indicates that placing sensors in the middle of the walkway and/or at less rigid locations of the floor structure is more desirable and tends to produce better results.

The results also show that the number of sensors can be significantly reduced in practical settings. Compared with using all four sensors for gait abnormality detection, there is a 1% to 7% performance drop when using one sensor only. This means that while combining multiple sensors boosts the detection accuracy, the improvement is marginal, and using a single sensor still offers better accuracy than the majority of the baseline model using four sensors (compare with Table III).

Moreover, integrating multiple sensor modalities may further enhance the performance by capturing multifaceted information. For example, existing studies using temporary setups of phone cameras provide complementary body motion information to our approach [60], which has been found to improve the accuracy for vibration-based gait analysis [35]. Wearable devices such as wrist band with Inertial measurement unit (IMU) [61] and electromyography (EMG) devices [4] provide extra information on upper extremity and muscle activation that are difficult to be captured by floor vibration. To this end, integrating these sensor modalities offers promising avenues for future research.

VI. CONCLUSION AND FUTURE WORK

In this study, we develop a new gait abnormality detection method using footstep-induced floor vibration for low-cost and non-intrusive in-home gait assessments. To achieve this, we develop a floor vibration sensing system and characterize how common gait abnormality types affect the resultant floor vibrations. The main challenge in developing our approach is high uncertainty in floor vibrations due to the different walking styles among people. To address this challenge, we develop a person-invariant feature transformation method through contrastive learning with a person- and gait-aware sampling strategy. This method learns an embedding space in which gait samples from different people stay close to each other while those from different gait types are far apart. After that, gait abnormalities are detected by comparing their embeddings with those of the normal gait. We design a visualization for personalized gait abnormality risk profiles to summarize the outcome of our method. We conducted a real-world walking experiment with 21 participants and obtained 85%, 90%, and 95% mean test accuracy in detecting gait abnormalities per gait cycle during initial contact, stance phase, and swing phase, respectively, and achieved a 0.81 F1 score in classifying four types of gait, including normal walking, gait with toe/flatfoot contacts, flexed-knee gait, and gait with foot drag. The evaluation results demonstrated the promise of floor vibration-based gait abnormality detection.

In the future, we aim to enable long-term, continuous tracking of gait changes by deploying our system in people's homes. To achieve this goal, we will focus on a specific type of gait disorder that is difficult to detect and collect data from real patients. During the data modeling development, we will address practical issues such as the data distribution disparities among various home environments and the integration of multiple sensing modalities. We plan to work with elder care centers to test our system and develop models to tackle challenges associated with multiple types of floors, sensors, and varying stages of diseases.

REFERENCE

[1] M. Pistacchi, M. Gioulis, F. Sanson, M. Ennio De Giovannini, G. Filippi, F. Rossetto, and S. Z. Marsala, "Gait analysis and clinical correlations in early parkinson's disease," *Functional neurology*, vol. 32, no. 1, p. 28, 2017.

[2] C. L. Martin, B. A. Phillips, T. Kilpatrick, H. Butzkueven, N. Tubridy, E. McDonald, and M. Galea, "Gait and balance impairment in early multiple sclerosis in the absence of clinical disability," *Multiple Sclerosis Journal*, vol. 12, no. 5, pp. 620–628, 2006.

[3] P. A. DELUCA, "Gait analysis in the treatment of the ambulatory child with cerebral palsy," *Clinical Orthopaedics and Related Research®*, vol. 264, pp. 65–75, 1991.

[4] Y. Guo, R. Gravina, X. Gu, G. Fortino, and G.-Z. Yang, "Emg-based abnormal gait detection and recognition," in *2020 IEEE International Conference on Human-Machine Systems (ICHMS)*. IEEE, 2020, pp. 1–6.

[5] V. Agostini, G. Balestra, and M. Knaflitz, "Segmentation and classification of gait cycles," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 5, pp. 946–952, 2013.

[6] J. Mickelborough, M. Van Der Linden, J. Richards, and A. Ennos, "Validity and reliability of a kinematic protocol for determining foot contact events," *Gait & Posture*, vol. 11, no. 1, pp. 32–37, 2000.

[7] Y. Guo, X. Gu, and G.-Z. Yang, "Mcdcd: Multi-source unsupervised domain adaptation for abnormal human gait detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 10, pp. 4017–4028, 2021.

[8] G. Cicirelli, D. Impedovo, V. Dentamaro, R. Marani, G. Pirlo, and T. R. D'Orazio, "Human gait analysis in neurodegenerative diseases: A review," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 229–242, 2021.

[9] D. Slijepcevic, M. Zeppelzauer, A.-M. Gorgas, C. Schwab, M. Schüller, A. Baca, C. Breiteneder, and B. Horsak, "Automatic classification of functional gait disorders," *IEEE journal of biomedical and health informatics*, vol. 22, no. 5, pp. 1653–1661, 2017.

[10] S. Potluri, S. Ravuri, C. Diedrich, and L. Schega, "Deep learning based gait abnormality detection using wearable sensor system," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 3613–3619.

[11] T. Liu, Y. Inoue, K. Shibata, and K. Shiojima, "A mobile force plate and three-dimensional motion analysis system for three-dimensional gait assessment," *IEEE Sensors Journal*, vol. 12, no. 5, pp. 1461–1467, 2011.

[12] B. Purcell, J. Channells, D. James, and R. Barrett, "Use of accelerometers for detecting foot-ground contact time during running," in *BioMEMS and nanotechnology II*, vol. 6036. SPIE, 2006, pp. 292–299.

[13] T.-N. Nguyen, H.-H. Huynh, and J. Meunier, "Skeleton-based abnormal gait detection," *Sensors*, vol. 16, no. 11, p. 1792, 2016.

[14] M. Chen, B. Huang, and Y. Xu, "Intelligent shoes for abnormal gait detection," in *2008 IEEE international conference on robotics and automation*. IEEE, 2008, pp. 2019–2024.

[15] J. Wu, B. Becsek, A. Schaer, H. Maurenbrecher, G. Chatzipirpiridis, O. Ergeneman, S. Pané, H. Torun, and B. J. Nelson, "Real-time gait phase detection on wearable devices for real-world free-living gait," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 3, pp. 1295–1306, 2022.

[16] M. Ullrich, N. Roth, A. Küderle, R. Richer, T. Gladow, H. Gaßner, F. Marxreiter, J. Klucken, B. M. Eskofier, and F. Kluge, "Fall risk prediction in parkinson's disease using real-world inertial sensor gait data," *IEEE journal of biomedical and health informatics*, vol. 27, no. 1, pp. 319–328, 2022.

[17] A. Elkholy, M. E. Hussein, W. Gomaa, D. Damen, and E. Saba, "Efficient and robust skeleton-based quality assessment and abnormality detection in human action performance," *IEEE journal of biomedical and health informatics*, vol. 24, no. 1, pp. 280–291, 2019.

[18] J. Fagert, M. Mirshekari, S. Pan, L. Lowes, M. Iammarino, P. Zhang, and H. Y. Noh, "Structure- and sampling-adaptive gait balance symmetry estimation using footstep-induced structural floor vibrations," *Journal of Engineering Mechanics*, vol. 147, no. 2, p. 04020151, 2021.

[19] Y. Dong, J. J. Zou, J. Liu, J. Fagert, M. Mirshekari, L. Lowes, M. Iammarino, P. Zhang, and H. Y. Noh, "Md-vibe: physics-informed analysis of patient-induced structural vibration data for monitoring gait health in individuals with muscular dystrophy," in *Adjunct proceedings of the 2020 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2020 ACM international symposium on wearable computers*, 2020, pp. 525–531.

[20] E. Kessler, V. V. S. Malladi, and P. A. Tarazaga, "Vibration-based gait analysis via instrumented buildings," *International Journal of Distributed Sensor Networks*, vol. 15, no. 10, p. 1550147719881608, 2019.

[21] Y. Dong and H. Y. Noh, "Structure-agnostic gait cycle segmentation for in-home gait health monitoring through footstep-induced structural vibrations," in *Society for Experimental Mechanics Annual Conference and Exposition*. Springer, 2023, pp. 65–74.

[22] S. Pan, M. Berges, J. Rodakowski, P. Zhang, and H. Y. Noh, "Fine-grained recognition of activities of daily living through structural vibration and electrical sensing," in *Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, 2019, pp. 149–158.

[23] Y. Dong, W. Yuyan, and H. Y. Noh, "Detecting gait abnormalities in foot-floor contacts during walking through footstep-induced structural vibrations," in *Proceedings of the 14th International Workshop on Structural Health Monitoring*, 2023.

[24] M. Lam, M. Mirshekari, S. Pan, P. Zhang, and H. Y. Noh, "Robust occupant detection through step-induced floor vibration by incorporating structural characteristics," in *Dynamics of Coupled Structures, Volume 4: Proceedings of the 34th IMAC, A Conference and Exposition on Structural Dynamics 2016*. Springer, 2016, pp. 357–367.

[25] S. Pan, N. Wang, Y. Qian, I. Velibeyoglu, H. Y. Noh, and P. Zhang, "Indoor person identification through footstep induced structural vibration," in *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications*, 2015, pp. 81–86.

[26] J. Fagert, M. Mirshekari, S. Pan, P. Zhang, and H. Y. Noh, "Gait health monitoring through footstep-induced floor vibrations," in *Proceedings of the 18th international conference on information processing in sensor networks*, 2019, pp. 319–320.

[27] B. T. Davis, B. I. Bryant, S. L. Fritz, R. Handlery, A. Flach, and V. A. Hirth, "Measuring gait parameters from structural vibrations," *Measurement*, vol. 195, p. 111076, 2022.

[28] J. M. Franco, Y. MejiaCruz, J. M. Caicedo, and Z. Jiang, "Feasibility of using accelerometers to detect human footsteps for cadence estimation on health sciences," in *Society for Experimental Mechanics Annual Conference and Exposition*. Springer, 2023, pp. 75–77.

[29] N. V. Boulgouris, D. Hatzinakos, and K. N. Plataniotis, "Gait recognition: a challenging signal processing technology for biometric identification," *IEEE signal processing magazine*, vol. 22, no. 6, pp. 78–90, 2005.

[30] S. Pan, T. Yu, M. Mirshekari, J. Fagert, A. Bonde, O. J. Mengshoel, H. Y. Noh, and P. Zhang, "Footprintid: Indoor pedestrian identification through ambient structural vibration sensing," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, pp. 1–31, 2017.

[31] Y. Dong, J. Fagert, and H. Y. Noh, "Characterizing the variability of footstep-induced structural vibrations for open-world person identification," *Mechanical Systems and Signal Processing*, vol. 204, p. 110756, 2023.

[32] J. Fagert, M. Mirshekari, P. Zhang, and H. Y. Noh, "Recursive sparse representation for identifying multiple concurrent occupants using floor vibration sensing," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 1, pp. 1–33, 2022.

[33] F. Li, J. Clemente, M. Valero, Z. Tse, S. Li, and W. Song, "Smart home monitoring system via footstep-induced vibrations," *IEEE Systems Journal*, vol. 14, no. 3, pp. 3383–3389, 2019.

[34] K. S. Hahm and B. W. Anthony, "In-home health monitoring using floor-based gait tracking," *Internet of Things*, vol. 19, p. 100541, 2022.

[35] Y. Dong, J. Liu, and H. Y. Noh, "Gaitvibe+ enhancing structural vibration-based footstep localization using temporary cameras for in-home gait analysis," in *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*, 2022, pp. 1168–1174.

[36] V. Racic, A. Pavic, and J. M. Brownjohn, "Experimental identification and analytical modelling of human walking forces: Literature review," *Journal of Sound and Vibration*, vol. 326, no. 1-2, pp. 1–49, 2009.

[37] J. Fagert, M. Mirshekari, S. Pan, L. Lowes, M. Iammarino, P. Zhang, and H. Y. Noh, "Structure-and sampling-adaptive gait balance symmetry estimation using footstep-induced structural floor vibrations," *Journal of Engineering Mechanics*, vol. 147, no. 2, p. 04020151, 2021.

[38] M. Mirshekari, J. Fagert, S. Pan, P. Zhang, and H. Y. Noh, "Step-level occupant detection across different structures through footstep-induced floor vibration using model transfer," *Journal of Engineering Mechanics*, vol. 146, no. 3, p. 04019137, 2020.

[39] Y. Dong, J. Zhu, and H. Y. Noh, "Re-vibe: Vibration-based indoor person re-identification through cross-structure optimal transport," in *Proceedings of the 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, 2022, pp. 348–352.

[40] Y. Dong, A. Bonde, J. R. Codling, A. Bannis, J. Cao, A. Macon, G. Rohrer, J. Miles, S. Sharma, T. Brown-Brandl *et al.*, "Pigsense: Structural vibration-based activity and health monitoring system for pigs," *ACM Transactions on Sensor Networks*, 2023.

[41] J. Chen, J. Benesty, Y. Huang, and S. Doclo, "New insights into the

noise reduction wiener filter," *IEEE Transactions on audio, speech, and language processing*, vol. 14, no. 4, pp. 1218–1234, 2006.
[42] R. Büssow, "An algorithm for the continuous morlet wavelet transform," *Mechanical Systems and Signal Processing*, vol. 21, no. 8, pp. 2970–2979, 2007.
[43] M. W. Whittle, *Gait analysis: an introduction*. Butterworth-Heinemann, 2014.
[44] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark *et al.*, "Learning transferable visual models from natural language supervision," in *International conference on machine learning*. PMLR, 2021, pp. 8748–8763.
[45] R. Girdhar, A. El-Nouby, Z. Liu, M. Singh, K. V. Alwala, A. Joulin, and I. Misra, "Imagebind: One embedding space to bind them all," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 15 180–15 190.
[46] A. v. d. Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," *arXiv preprint arXiv:1807.03748*, 2018.
[47] B. Tao, L. Huang, H. Zhao, G. Li, and X. Tong, "A time sequence images matching method based on the siamese network," *Sensors*, vol. 21, no. 17, p. 5900, 2021.
[48] C. Zhang, K. Zhang, C. Zhang, T. X. Pham, C. D. Yoo, and I. S. Kweon, "How does simsiam avoid collapse without negative samples? a unified understanding with self-supervised contrastive learning," *arXiv preprint arXiv:2203.16262*, 2022.
[49] M. Kang and B. Lee, "Tictok: Time-series anomaly detection with contrastive tokenization," *IEEE Access*, 2023.
[50] R. Zhang, P. Isola, and A. A. Efros, "Colorful image colorization," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14*. Springer, 2016, pp. 649–666.
[51] P. Bachman, R. D. Hjelm, and W. Buchwalter, "Learning representations by maximizing mutual information across views," *Advances in neural information processing systems*, vol. 32, 2019.
[52] A. Kolesnikov, X. Zhai, and L. Beyer, "Revisiting self-supervised visual representation learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 1920–1929.
[53] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *International conference on machine learning*. PMLR, 2020, pp. 1597–1607.
[54] W. S. Lee and B. Liu, "Learning with positive and unlabeled examples using weighted logistic regression," in *ICML*, vol. 3, no. 2003, 2003, pp. 448–455.
[55] I. Melekhov, J. Kannala, and E. Rahtu, "Siamese network features for image matching," in *2016 23rd international conference on pattern recognition (ICPR)*. IEEE, 2016, pp. 378–383.
[56] M. Shorfuzzaman and M. S. Hossain, "Metacovid: A siamese neural network framework with contrastive loss for n-shot diagnosis of covid-19 patients," *Pattern recognition*, vol. 113, p. 107700, 2021.
[57] Vicon motion capture system. [Online]. Available: https://www.vicon.com/applications/life-sciences/
[58] Y. Dong, Y. Wu, J. R. Codling, J. Aggarwal, P. Huang, W. Ding, H. Latapie, P. Zhang, and H. Y. Noh, "Gamevibes: Vibration-based crowd monitoring for sports games through audience-game-facility association modeling," in *Proceedings of the 10th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, 2023, pp. 177–188.
[59] Y. Dong, M. Iammarino, J. Liu, J. Codling, J. Fagert, M. Mirshekari, L. Lowes, P. Zhang, and H. Noh, "Ambient floor vibration sensing advances accessibility of functional gait assessment for children with muscular dystrophies," 2023.
[60] S. D. Uhlrich, A. Falisse, Ł. Kidziński, J. Muccini, M. Ko, A. S. Chaudhari, J. L. Hicks, and S. L. Delp, "Opencap: Human movement dynamics from smartphone videos," *PLoS computational biology*, vol. 19, no. 10, p. e1011462, 2023.
[61] S. Cerfoglio, P. Capodaglio, P. Rossi, I. Conforti, V. D'Angeli, E. Milani, M. Galli, and V. Cimolin, "Evaluation of upper body and lower limbs kinematics through an imu-based medical system: A comparative study with the optoelectronic system," *Sensors*, vol. 23, no. 13, p. 6156, 2023.

ACKNOWLEDGEMENT

This work was funded by the U.S. National Science Foundation (under grant numbers NSF-CMMI-2026699) and Stanford Blume Fellowship. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either express or implied, of any University, the National Science Foundation, or the United States Government or any of its agencies.