

Pushing the Limits of WiFi-based Gait Recognition Towards Non-gait Human Behaviors

Dawei Yan, Panlong Yang, *Senior Member, IEEE*, Fei Shang, Feiyu Han,
Yubo Yan, *Member, IEEE*, and Xiang-Yang Li, *Fellow, IEEE*

Abstract—WiFi-based gait recognition technologies have seen significant advancements in recent years. However, most existing approaches rely on a critical assumption: users must walk continuously and maintain a consistent body posture. This poses a substantial challenge when users engage in non-periodic or discontinuous behaviors (*e.g.*, stopping, starting, or turning mid-walk), which can disrupt the extraction of gait-related features and degrade recognition performance. To address this issue, we propose *freeGait*, a novel approach designed to mitigate the impact of non-gait behaviors in WiFi-based gait recognition systems. Our solution models this problem as domain adaptation, where we learn domain-independent representations to isolate gait features from behavior-dependent noise. We treat human behaviors with labeled user data as source domains and behaviors without user labels as target domains. However, applying domain adaptation directly is challenging due to the ambiguous classification boundaries in the target domains for WiFi signals. To overcome this, we align the posterior distributions between the source and target domains and constrain the conditional distribution within the target domains to enhance gait classification accuracy. Additionally, we implement a data augmentation module to generate data resembling the labeled data, while supervised learning ensures distinctiveness between users. Our experiments, conducted with 20 participants across 3 different scenarios, demonstrate that *freeGait* can accurately predict data across 15 domains by labeling only a small subset from 6 source domains, achieving up to a 45% improvement in user classification accuracy compared to existing methods.

Index Terms—WiFi-based Sensing, Gait Recognition, Data Augmentation, Domain Adaptive Network.

I. INTRODUCTION

HUMAN identification has become a crucial factor in various applications, including security management and personalized services in smart spaces [1]–[3]. Numerous studies have explored key aspects of human identification using different technologies, but they all face significant limitations. For example, video-based methods [4], [5] are sensitive to lighting conditions and raise privacy concerns. Acoustic signal-based methods [6], [7] are vulnerable to environmental noise and mimicry attacks. Wearable device-based methods [8], [9] require users to wear devices, leading to discomfort. Millimeter-wave radar-based methods [10], [11] entail additional hardware deployment costs and have limited

D. Yan, F. Shang, Y. Yan and X. Li are with School of Computer Science and Technology, University of Science and Technology of China, Hefei 230021, China. (e-mail: yandw@mail.ustc.edu.cn, shf_1998@outlook.com, yuboyan@ustc.edu.cn, xiangyangli@ustc.edu.cn).

P. Yang and F. Han are with School of Computer Science, Nanjing University of Information Science & Technology, Nanjing 210044, China. (e-mail: plyang@ustc.edu.cn, fyhan@mail.ustc.edu.cn).

sensing ranges. Biometric methods, such as facial recognition, fingerprint, and iris scanning [12], [13], are associated with privacy issues, contact requirements, and high costs. Furthermore, infrared camera-based methods [14]–[16] are easily affected by occlusion and temperature, and require additional installation and maintenance investments for large-scale deployment. While they address privacy concerns associated with RGB images, infrared cameras can still cause user discomfort due to the perception of surveillance.

Compared to these advanced technologies, WiFi-based human recognition, particularly gait recognition, has gained significant attention recently due to its advantages of ubiquity, non-contact, and non-invasion [1], [2], [17]–[19]. The basic principle is that as a person walks, their movements disrupt WiFi signals, such as *channel state information* (CSI), and each individual's walking pattern is distinct. These differences in limb movement and speed serve as a unique signature of a person's gait, enabling user identification. While previous WiFi-based gait recognition systems have offered certain conveniences for human identification, significant challenges remain in deploying these systems in the real world.

In particular, current WiFi-based gait recognition systems are typically developed under the assumption that the walking patterns of users are periodic. The effectiveness of CSI in capturing gait patterns lies in the physical principles of multipath signal propagation. When a person walks, their periodic limb movements modulate the wireless signals traveling between the transmitter and receiver. These movements introduce predictable changes in the amplitude and phase of the CSI signals, forming distinctive patterns. Specifically, the rhythmic swinging of arms and legs creates consistent fluctuations in multipath signal components, which manifest as periodic variations in the time-frequency domain. However, we observe that in real-world scenarios, users not only move their legs when walking, but may also be accompanied by various other non-periodic human behaviors (*e.g.*, turning around, bending, and carrying luggage), as shown in Fig. 1. What's more, in many cases, people often do not walk continuously; instead, they may start, stop, and change pace. Such non-periodicity and discontinuity are challenging for scenarios that require long-term human recognition, as the system may fail to comprehensively capture the distinctive gait-related features, leading to performance issues. Therefore, in this paper, we raise a problem: “*Could WiFi technology have the potential to extract stable gait features accompanying various non-gait human behaviors?*”

Most WiFi-based gait recognition schemes rely on CSI to

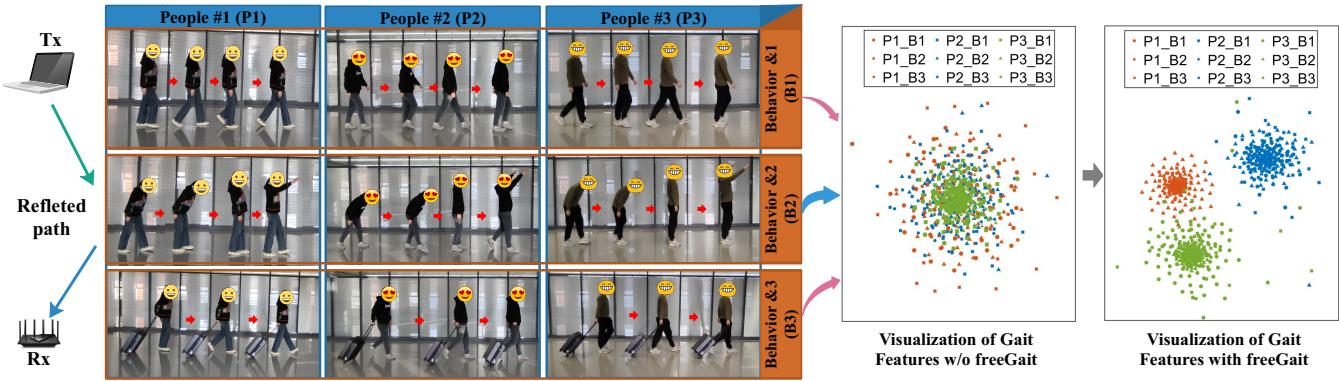


Fig. 1. Users are often accompanied by various gait or non-gait human behaviors when walking, e.g., normal continuous walking (B1), bending or waving (B2), carrying luggage (B3). WiFi propagation signals under different non-gait human behaviors are different, which leads to inconsistency in user gait patterns based on WiFi, thus making the gait recognition system ineffective. *freeGait* can accurately extract the users' gait features despite their accompanying various non-gait behaviors.

extract features related to gait. Since CSI captures detailed channel characteristics, it is highly sensitive to variations in human behaviors. As shown in Fig. 1, we give an example. It can be seen that the reflection paths between the laptop (*i.e.*, Tx) and the WiFi router (*i.e.*, Rx) change due to differences in human behaviors, resulting in the inability to accurately separate gait features even if the same people walks along the same path, and ultimately making different users' gait features are mingled together. It indicates that human non-gait patterns extremely interfere with the distribution of gait features. A *more detailed analysis of human behaviors and gait patterns is given in Section II-B*.

Theoretically, if we could gather sufficient CSI data capturing every possible behavior of each user, we could train a gait recognition model resilient to variations in human behavior. However, given the vast diversity and complexity of human behaviors, as well as the influence of different walking paths on gait patterns, this approach would be exceedingly labor-intensive. Moreover, many non-gait behaviors and walking paths might not be accounted for, making this solution impractical in real-world applications. An alternative approach could involve integrating human behavior identification with gait recognition systems, allowing for the accurate detection and separation of human behaviors from the mixed CSI data. However, this still requires comprehensive data collection for all possible behaviors, and signal separation remains a challenging task.

Furthermore, although some studies have attempted to mitigate the impact of walking paths by estimating the walking direction, this typically requires setting up multiple transceivers and ensuring that users walk normally and continuously within a specific area [2], [20], [21]. However, such setups are cumbersome for many scenarios and users, limiting the practicality of gait recognition systems. In addition, existing methods predominantly focus on recognizing authorized users within collected gait data, while seldom addressing the detection of abnormal individuals. Especially, user behavior diversity and environmental noise impact the accuracy of anomaly detection. The diversity in user behaviors increases the risk of misclassification, as anomalous behaviors can overlap with normal gait patterns, and unstable CSI quality can significantly

reduce detection precision. Due to these limitations, current approaches are not only impractical for real-world applications but also difficult to scale across broader scenarios. Therefore, in this paper, we focus on achieving accurate gait recognition for users engaged in various human behaviors using only a single pair of WiFi transceivers, while also enabling real-time detection of anomalous users to enhance the system's security and adaptability.

Therefore, to develop a behavior-robust gait recognition system with only a pair of WiFi transceivers, we face three key challenges:

- Users are often accompanied by various non-periodic or discontinuous human behaviors while walking, each exerting a different influence on WiFi-based gait patterns. To eliminate these influences, we would need to account for all possible human behaviors during gait data collection to accurately capture the user's identity. However, this is challenging to achieve.
- Other items, such as walking paths and speeds, also affect the gait patterns of users. To address these influences, we would need to account for all possible variations in these factors, even when the same human behavior is performed. However, it is impractical to label data for every possible combination of these factors associated with the user.
- Extracting fine-grained gait patterns using a pair of WiFi transceivers is challenging. When the user is positioned far from the WiFi transceiver or walking in a direction nearly parallel to it, the CSI noise can overwhelm and obscure the gait pattern. Besides, the gait patterns of abnormal users are unseen and easily confused with those of legal users, increasing the difficulty of detection.

To address the above challenges, we present *freeGait*, a WiFi gait recognition system based on data processing methods and a fine-grained deep learning framework, only by collecting and labeling a small amount of CSI data from one WiFi transceiver.

Firstly, we perform a series of processing on the raw CSI and obtain refined spectrograms reflecting the users' gait patterns. Specifically, we observe that using *principal component analysis* (PCA) at the subcarrier level and extracting the first

principal component can better remove environmental noise. Then, to eliminate *automatic gain control* (AGC) noise, while retaining the original data characteristics, we use *density-based spatial clustering of applications with noise* (DBSCAN) on the CSI and extract the class with the highest density. These techniques enable the effective extraction of dynamic target signals, even when the user is positioned far from the transceiver. Additionally, we use *short-time Fourier transform* (STFT) to obtain the spectrograms.

Secondly, we model the acquisition of gait features independent of human behaviors and walking paths as the domain adaptation problem. Specifically, *freeGait* considers human behaviors and walking paths in the database with users' IDs as the source domains, and unknown human behaviors and walking paths without users' IDs as the target domains. The network model is then jointly trained using labeled user's ID data in the source domains and unlabeled data in the target domains, and the ultimate goal is to learn common features (gait patterns) of the labeled data and unlabeled data, while downplaying the differences between source and target domains (effects of different human behaviors and walking paths). In this way, *freeGait* can predict users from different human behaviors and walking paths without relabeling users' IDs for new data.

Although there has been some domain adaptation works in WiFi sensing [22]–[27], they mainly focus on tasks such as gesture recognition, presence detection, localization, and targeting challenges associated with environmental or device settings changes. While these approaches achieve robustness in their respective tasks, they do not address the unique challenge of behavioral diversity in gait recognition. Our work fills this gap by modeling the acquisition of gait features independent of human behavior and walking paths as a domain adaptation problem. The core idea of using domain adaptation in gait recognition is to align feature distributions across different domains (*e.g.*, behaviors and walking paths) while preserving gait-related physical characteristics, such as the periodicity of human motion and Doppler shifts introduced by limb movement. In domain adaptation, we aim to align the periodic gait features and Doppler shifts across different domains. This alignment helps to reduce the variation caused by changes in human behaviors and walking paths. Specifically, adversarial domain adaptation and feature alignment techniques are used to map the gait features from different domains into a shared space, where gait-related patterns are preserved, and non-gait related variations are minimized.

Thirdly, to obtain enough source domain data to eliminate the influence of different human behaviors and other items, we utilize data augmentation technology to generate synthetic data similar to the collected labeled data with users' IDs, allowing us to scale the labeled data to cover a wide range of human behaviors and other items. Specifically, we collect some data for each user walking along different paths (*e.g.*, one-minute data), combine it with human behaviors data in the source domains, and use an *adversarial autoencoder* (AAE) [28], generating similar but different gait data for each user separately. In addition, we incorporate the idea of supervised learning to avoid generating extremely close sample data among different

users. In this way, *freeGait* can predict users' gait under more human behaviors, and walking paths.

Finally, we also detect intruders by identifying unauthorized users. Specifically, we achieve this by setting a threshold based on class density and comparing the distance between a new user and the nearest neighbor class to determine whether the user is unauthorized.

In summary, the main contributions of this paper are:

- In this paper, we analyze the impact of users' non-periodic or discontinuous behaviors on WiFi-based gait recognition. Then, we propose *freeGait*, a WiFi-based gait recognition system, that aims to mitigate users' diverse non-gait behaviors while maintaining accurate gait recognition.
- We design domain adaptation techniques to reduce the coupling of gait patterns with behaviors and paths, and enable *freeGait* to learn behavior- and path-independent gait features. Then, we design a data augmentation method so that using only a small amount of labeled data with users' IDs, *freeGait* can perform well in processing a variety of rich gait data of behaviors and paths without user IDs.
- We implement *freeGait* with a pair of WiFi transceivers and conduct extensive experiments. We define 6 different human behaviors, 3 different walking paths, and 3 different walking speeds, and recruit twenty volunteers of different heights and weights to participate in the experiment. The results show that accurate predictions for a total of 15 domains can be achieved by collecting and labeling only a small amount of data from 6 source domains, which is at worst 45% improvement over other existing techniques.

The rest of this paper is organized as follows: Section II presents the basic principles of gait recognition based on WiFi CSI and the impact of human behavior on gait features. We show the overview of *freeGait* in Section III. In Section IV, we describe the module design of the proposed framework in detail. Implementation and evaluation are presented in Section V and Section VI. In Section VII, we discuss the limitations of this work and future work. Section VIII discusses the related works. We finally conclude our work in Section IX.

II. PRELIMINARIES

A. Gait Recognition Based on WiFi CSI

Channel State Information (CSI) captures the intricate characteristics of WiFi signal propagation, detailing various aspects of the transmission path, including signal attenuation, multipath effects, and scattering. Specifically, CSI can be represented as the values of each element within the channel matrix \mathbf{H} [29]. Modern WiFi technologies often employ *Orthogonal Frequency Division Multiplexing* (OFDM) and *Multiple Input Multiple Output* (MIMO), where each channel between transmitter and receiver pairs (TX-RX) is composed of multiple subcarriers across multiple antenna pairs. This inherently enriches the data available for CSI-based sensing solutions. In particular, some work has developed tools that can extract CSI from received WiFi packets and obtain the

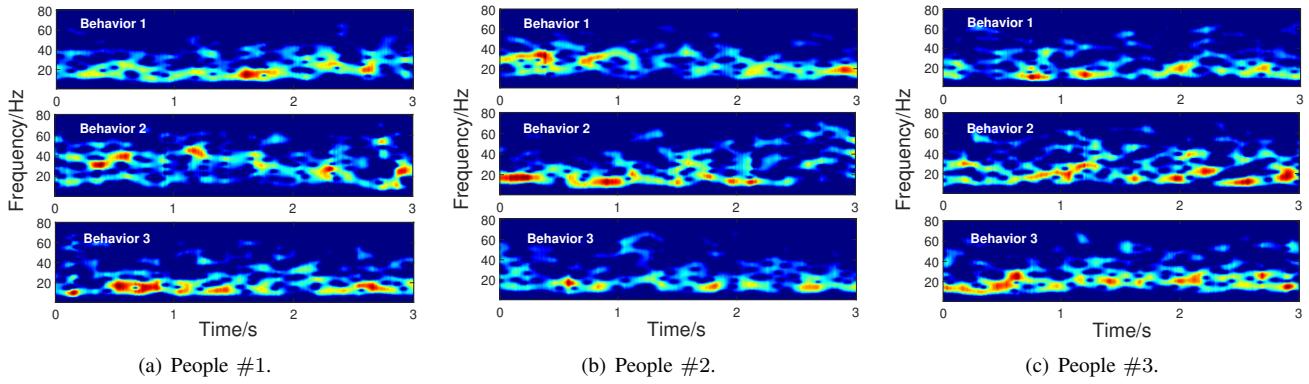


Fig. 2. Spectrograms of three different people walking with three different human behaviors (behavior 1: normal continuous walking; behavior 2: stop-and-go and bending; behavior 3: individual carrying luggage). Even for the same people, gait patterns are different in different human behaviors.

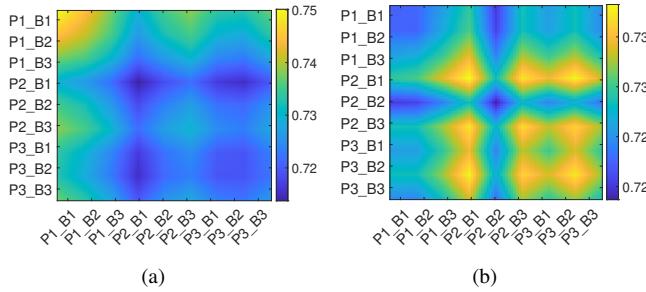


Fig. 3. Structural Similarity Index Measure (SSIM) of the Spectrograms of the same or different people and behaviors with different walking paths: (a) tr1 and (b) tr2.

channel matrix \mathbf{H} , such as *Intel 5300 CSI Tool* [30], *Atheros CSI Tool* [31] and *Nexmon CSI Extractor* [32], which has made WiFi sensing solutions based on CSI flourish in recent years.

Similar to other CSI-based passive WiFi sensing systems, CSI-based gait recognition leverages the fundamental principle that human movement induces distinct disturbances in WiFi signals. When individuals with varying walking patterns move within a fixed Tx-Rx setup, they either absorb or reflect WiFi signals in unique ways, leading to distinct variations in the channel matrix \mathbf{H} [2], [17]. Given that the CSI phase can be influenced by numerous factors, this paper focuses exclusively on using the amplitude of CSI as input:

$$\mathbf{H}[t] = (|h_{1,1}[t]|, \dots, |h_{1,n_{sc}}[t]|, \dots, |h_{n_{ss},n_{sc}}[t]|), \quad (1)$$

where $h_{i,j}[t]$ represents the CSI value of the i -th stream on the j -th subcarrier collected at time t , n_{sc} and n_{ss} are the numbers of spatial streams and frequency subcarriers, and $|\cdot|$ denotes the amplitudes of complex numbers. Then we use the preprocessed input parameters (*e.g.*, sub-spectrograms) to extract gait features related to individuals. Detailed processing is given in Section IV-A.

B. Human Behaviors Impact Analysis

When users walk at different times and engage in various behaviors, the WiFi signals reflected by their bodies differ due to complex and dynamic multipath effects, even for the same individual. For continuous gait patterns, the CSI signal exhibits periodic characteristics due to the regular movement of the human body, particularly the limbs. The regularity of

gait movements results in periodic changes in signal amplitude and phase, which can be captured and used to extract gait features. These features are typically smooth, with consistent frequency and amplitude patterns that reflect the cyclic nature of human walking. In contrast, non-periodic behaviors lead to discontinuous or irregular CSI patterns. These behaviors result in sharp, erratic changes in signal amplitude and phase, which are not characteristic of continuous walking. For example, when a person stops or changes direction abruptly, the CSI signal may exhibit sudden drops, spikes, or rapid fluctuations that disrupt the otherwise smooth, periodic signal patterns associated with gait.

To examine how different human behaviors impact CSI, we conduct preliminary tests in a $12m \times 10m$ space. We invite three volunteers (*i.e.*, P1, P2, and P3) with varying heights and weights to participate. Each volunteer walks along the same path within a designated area and performs three distinct behaviors: normal continuous walking (B1), stop-and-go and bending (B2), and individual carrying luggage (B3), as depicted in Fig. 1. The Tx and Rx are established using an *Industrial Personal Computer* (IPC) equipped with an Intel 5300 Network Interface Card (NIC), operating at the 5GHz/HT40- ISM band, with a packet sending rate set to 1000Hz. CSI is collected for each volunteer across the different behaviors, with each condition repeated three times. Notably, our analysis focuses solely on the influence of human behaviors on gait patterns, ensuring that volunteers walk the same path each time, with the scenario (including the environment and transceiver setup) remaining unchanged.

We extract the amplitude of each CSI packet and focus on analyzing the gait patterns of different individuals using the 30 subcarriers from the first antenna of the Rx. The raw CSI amplitude undergoes a series of preprocessing steps, including data denoising and Short-Time Fourier Transform (STFT), to generate spectrograms for each condition, as detailed in Section IV-A. Fig. 2 illustrates the gait patterns of three volunteers as they walk while performing three distinct human behaviors. Even when the same individuals walk along the same path and in the same direction, the accompanying behaviors lead to noticeable variations in gait patterns.

In fact, these behavior-induced variations can sometimes exceed the differences between the gait patterns of different

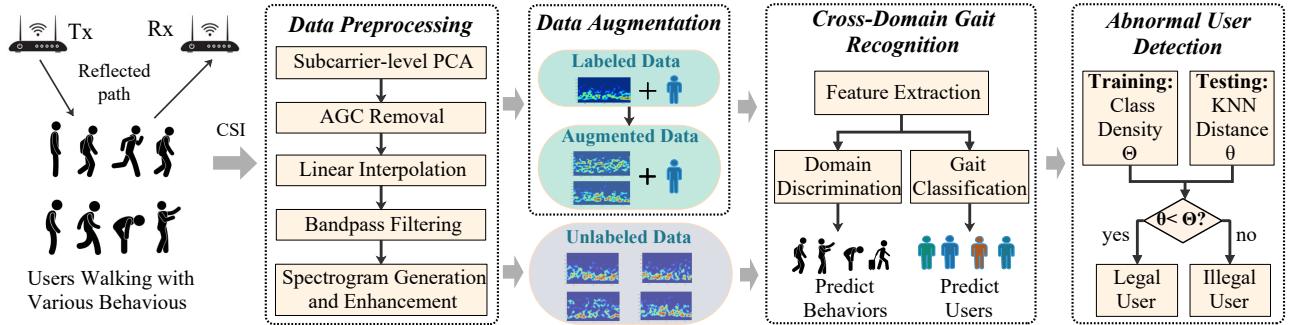


Fig. 4. Overview of *freeGait*: consists of data preprocessing, data augmentation, cross-domain gait recognition, and abnormal user detection.

individuals. To analyze the impact of human behaviors on gait patterns in detail, we calculate the *Structural Similarity Index Measure* (SSIM) of the spectrograms under the same or different conditions, and the results are shown in Fig. 3, where the three users and three behaviors in Fig. 3(a) and Fig. 3(b) are the same, but with different walking paths (tr1 and tr2). Intuitively, we can see that human behavior confuses the correlation of user gait patterns and makes them difficult to distinguish. Consequently, gait recognition accuracy may decline when subjects engage in untrained behaviors during walking. This analysis underscores the significant impact of accompanying human behaviors on gait patterns. We then demonstrate how our proposed techniques and models mitigate this effect, relying on only a small amount of labeled gait data from different behaviors.

III. OVERVIEW

A. Problem Definition

The goal of this paper is to make the WiFi-based gait recognition system independent of the subject's human behaviors and walking paths, and only by labeling a small amount of gait sample data during one or a few behaviors, the subject's gait can be accurately identified, even if accompanied by more diverse behaviors or paths. As shown in Fig. 4, we place one Tx and one Rx in the physical space to form a WiFi propagation link, so that a gait sensing area can be constructed. Specifically, we collect the CSI data of each subject when walking within the sensing area and input them into *freeGait*. Among them, a small part of the data (*i.e.*, known behaviors and paths) carries users' IDs, and the remaining data (*i.e.*, unknown behaviors and paths) does not carry users' IDs. The final result should be one that can accurately predict the users' IDs corresponding to all gait data (*i.e.*, labeled and unlabeled data with different behaviors and paths).

B. *freeGait*'s Architecture

As shown in Fig. 4, we place one Tx and one Rx in the physical space to form a WiFi propagation link, so that a gait sensing area can be constructed. *freeGait* mainly consists of five modules: data preprocessing, data augmentation, feature extraction, gait classification, and domain discrimination.

Data preprocessing. We extract the amplitude from the original CSI data collected. Then, we perform PCA at the subcarrier level to filter out environmental noise and remove

the influence of AGC to obtain clean CSI. Next, we perform linear interpolation and bandpass filtering to obtain the equally sampled gait-related signals of interest to facilitate the extraction of gait patterns. Finally, we perform STFT on each piece of CSI data to generate and enhance the spectrogram as input to the model.

Data augmentation. A small part of the data (*i.e.*, known behaviors and paths) with users' IDs enters the AAE-based data augmentation, and each user's labeled data enters an AAE to generate data separately to expand more potential human behaviors and walking paths.

Cross-domain gait recognition. The augmented labeled data (source domains data) and unlabeled data (target domains data) are input the cross-domain gait recognition to learn gait features independent of non-gait human behaviors and walking paths.

Abnormal user detection. During training, we calculate the average neighbor distance of each legitimate category as the category density Θ . During testing, we identify the average distance θ of the *K-nearest neighbors* (KNN) of the test sample, and if it is greater than the weighted category density, the test sample is considered to be from an illegal user.

IV. SYSTEM DESIGN

A. Data Preprocessing

Given the raw CSIs, we perform a series of preprocessing on them to obtain many refined spectrograms that can be used to extract gait features. In this paper, we use the Linux 802.11n CSI Tool [30] to collect CSIs based on the Intel 5300 NIC. Tx is equipped with one antenna and Rx is equipped with three antennas. We set the sampling rate to 1000Hz, and each CSI packet contains 30 OFDM subcarriers. Thus, we can get about 90000 CSI values per second.

Subcarrier-level PCA. Raw CSI collected using WiFi NICs contains significant noise [33], primarily manifesting as environmental noise and AGC noise [34], [35]. Environmental noise typically follows a Gaussian distribution throughout the space [36]. In our setup, we use a single WiFi transceiver to continuously recognize the user's gait during the entire walking process. When the user is distant from the transceiver or walks in a direction nearly parallel to the propagation path, the gait pattern can be overwhelmed by noise. Traditional denoising methods (*e.g.*, lowpass, wavelet) tend to perform poorly under these conditions [37]. Moreover, the effectiveness

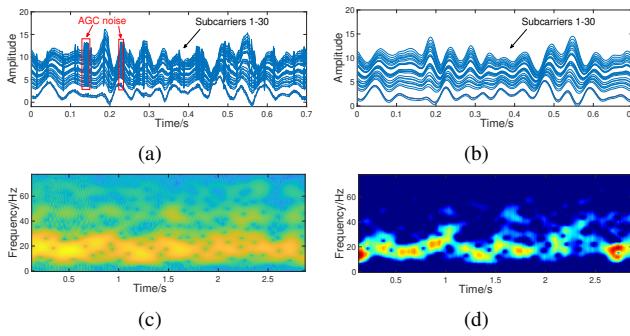


Fig. 5. Data preprocessing: (a) CSI after subcarrier-level PCA. (b) CSI after AGC removal. (c) Raw spectrogram. (d) Enhancement spectrogram.

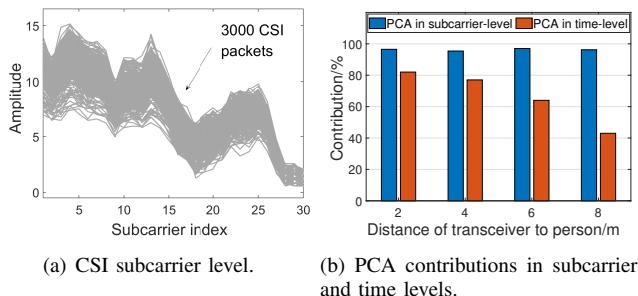


Fig. 6. The first component of the subcarrier-level PCA is higher than time-level PCA.

of conventional PCA denoising in the time dimension hinges on the selection of the number of principal components; choosing too many or too few components may either retain environmental noise or lose important target signals. When dynamic signals are masked by environmental noise, PCA becomes less effective as the principal components do not represent the dynamic signal.

To address this, we shift our focus to the subcarrier level. By transposing the CSI matrix, we represent the reflection paths of all subcarriers in the environment over time. As illustrated in Fig. 6(a), at the subcarrier level, changes in CSI amplitude correspond to alterations in reflection paths caused by human motion. Assuming a static environment, CSI data in the subcarrier dimension exhibit a higher correlation. We employ a covariance method to compute PCA on CSI data collected at various distances and measure their correlation [38], [39]. As shown in Fig. 6(b), while the contribution of the first principal component in the time dimension decreases with increasing distance, the contribution of the first principal component in the subcarrier dimension remains consistently high across different distances. Thus, by applying PCA at the subcarrier level, we can effectively remove environmental noise using just the first principal component, even when the noise significantly masks the gait pattern, as demonstrated in Fig. 5(a).

AGC removal. After removing environmental noise, we also need to address the influence of AGC noise. Previous approaches have effectively removed AGC noise using the ratio of two antennas [40]. However, in this paper, we avoid the ratio method to fully preserve the original distribution of CSI data. Our observations reveal that AGC noise manifests as uncertain and sparse points surrounding the dynamic signal in the time dimension [41], [42]. We can filter out this noise

by targeting its sparse density distribution. Specifically, we employ the DBSCAN spatial clustering method [43] to cluster AGC-related points based on their sparse distribution density and retain only the cluster with the highest number of scatter points [39]. The effectiveness of this approach is illustrated in Fig. 5(b).

Linear interpolation and bandpass filtering. During the process of collecting CSI by WiFi NIC, a small number of data packets will not be evenly distributed over time. However, we execute STFT based on a fixed time window (3 seconds) and generate a spectrogram as an initial feature input to the network model. Therefore, we linearly interpolate the data of each subcarrier at equal time intervals based on the timestamp of the CSI packet and according to the sampling rate of 1000Hz. In addition, the effective frequency range related to gait is from 10Hz to 70Hz [20], so we use a Butterworth-based bandpass filter to remove the signals outside the effective frequency range.

Spectrogram generation and enhancement. To represent the gait features of different people while walking, we perform STFT on the CSI processed above to obtain a spectrogram reflecting time and frequency information. Specifically, we first apply a sliding window to slice the time series CSI, each segment contains 3s (*i.e.*, 3000 CSI packets), and the sliding window is 1s (*i.e.*, 1000 CSI packets apart). Then, we perform STFT on each slice. To maintain the balance of time and frequency resolution, we set the FFT size to 1024 and the sliding window step size to 6, which can achieve a frequency resolution of 0.95Hz and a time resolution of 6 ms, which has a better discrimination effect on human gait signals between 10Hz and 70Hz. The spectrogram is shown in Fig. 5(c). We only show the image of 0-80Hz, where yellow indicates higher reflected energy.

Furthermore, to reduce the noise of the spectrogram to obtain a refined spectrogram, we used some technologies in WiFiU [17] to enhance each spectrogram. Specifically, we add the amplitudes of the corresponding spectrograms of 30 subcarriers to obtain the superimposed spectrogram, and only retain 0-80Hz. We then normalize each FFT block and subtract the mean of the amplitude of the entire spectrogram to remove background noise (anything less than 0 is set to 0). Finally, we apply a two-dimensional Gaussian filter with size=10 and $\delta = 0.4$ to obtain the enhanced spectrogram, and the result is shown in Fig. 5(d).

B. Cross-Domain Gait Recognition

The deep learning network framework for cross-domain gait recognition in *freeGait* is shown in Fig. 7. It mainly consists of feature extraction, gait classification, and domain discrimination. Specifically, we input the augmented and unlabeled data into the feature extraction module based on *convolutional neural network* (CNN) and *long short-term memory* (LSTM) to generate latent features for low-dimensional representation. Then, based on the generated latent features, the gait classification module is used to obtain the predicted user with maximized classification accuracy. At the same time, to eliminate the influence of different human behaviors and

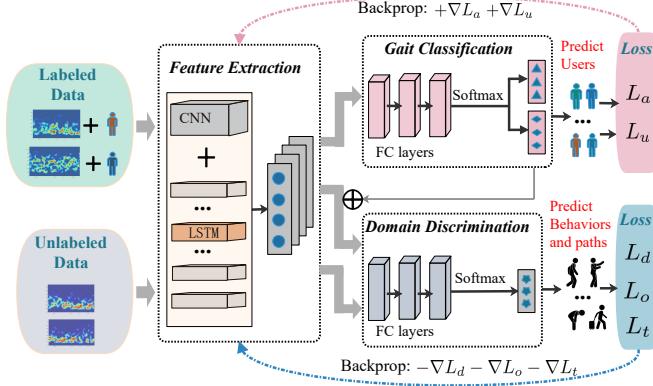


Fig. 7. The framework for learning behavior- and path-independent gait features using domain adaptation is used in *freeGait*.

walking paths (*i.e.*, domain-specific features), we design a domain discrimination module to predict each domain. The goal of the domain discrimination module is to maximize the accuracy of domain labeling, which seems to contradict learning domain-independent features. However, the feature extraction module is designed to try its best to deceive the domain discrimination module, *i.e.*, to minimize the prediction accuracy of behaviors and paths, while improving the user gait prediction accuracy. In this way, we achieve learning common features that are independent of human behaviors and walking paths that define users.

Feature extraction. As shown in Fig. 7, we input all labeled and unlabeled data together into the feature extractor to output their feature vectors. We use the widely adopted CNN and LSTM deep learning architectures to extract gait features [44], [45]. Particularly, we use three-layer stacked CNN and three-layer stacked LSTM. As shown in Fig. 8, at each layer of the CNN, we use convolutional layers with 2D convolution kernels, utilize batch normalization layers to speed up training with the batch size is 256, insert *rectified linear units* (ReLUs) to introduce nonlinearity, and use max pooling layers to reduce the size of the representation. In addition, LSTM has good performance in time series data processing. It is used to learn the temporal dynamic features extracted by CNN. Each LSTM layer has the same number of neurons and the number of each LSTM is 128, and uses the Sigmoid activation function. Therefore, given the input data \mathbf{S}_i , we can obtain the features \mathbf{Z}_i through CNN and LSTM:

$$\mathbf{Z}_i = \text{CNN}(\mathbf{S}_i; \Theta_{cnn}) \oplus \text{LSTM}(\mathbf{S}_i; \Theta_{lstm}), \quad (2)$$

where Θ_{lstm} and Θ_{cnn} are the parameters of CNN and LSTM, \oplus represents the operation of concatenation.

Gait classification. As shown in Fig. 7, after obtaining the feature \mathbf{Z}_i , we use three fully connected layers [46] and the activation function ReLU to learn the representation \mathbf{V}_i of \mathbf{S}_i , and let \mathbf{V}_i through an output layer with an activation function of softmax to obtain the predicted probability vector \hat{y}_i of the gait. It is worth noting that the reason why three fully connected layers are used is to obtain more parameters, and more fully connected layers do not improve the performance much. In addition, to improve the accuracy of gait classification, we

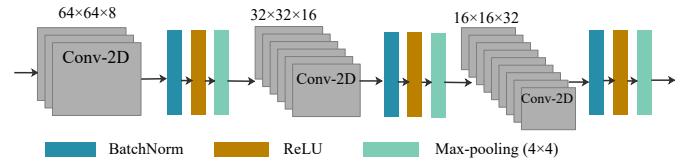


Fig. 8. Three-layer CNN architecture for feature extraction, each CNN layer uses 2D convolution with batch normalization, ReLU activation for nonlinearity, and max pooling to reduce representation size.

use a combination of supervised learning and unsupervised learning [47]. Specifically, we predict user labels \hat{y}_i^a and \hat{y}_i^u for labeled and unlabeled data respectively. For all data, we use cross-entropy as the loss function for gait classification:

$$\begin{aligned} \mathcal{L}_a &= -\frac{1}{n_a} \sum_{i=1}^{n_a} \sum_{k=1}^K y_{ik}^a \log (\hat{y}_{ik}^a), \\ \mathcal{L}_u &= -\frac{1}{n_u} \sum_{i=1}^{n_u} \sum_{k=1}^K \hat{y}_{ik}^u \log (\hat{y}_{ik}^u), \end{aligned} \quad (3)$$

where n_a and n_u are the numbers of labeled and unlabeled data used for training, and K is the total number of users.

Domain discrimination. *Domain-adversarial training of neural networks* (DANN) is a special case of transfer learning [48]. We use the idea of DANN to eliminate the effects of human behaviors and walking paths. Specifically, we define different walking paths and human behaviors as different domains, and the domain discriminator is used to identify different walking paths and human behaviors. Our goal is to enable the feature extractor to fool the domain discriminator, thereby producing gait features that are independent of human behaviors and walking paths.

As shown in Fig. 7, we input the output \mathbf{Z}_i of the feature extractor into the domain discriminator and predict the domain label \hat{d}_i through the same process. The domain discriminator is also composed of three fully connected layers with the activation function ReLU and an output layer with the softmax activation function. We use cross-entropy as the loss function for domain label prediction:

$$\mathcal{L}_d = -\frac{1}{n_d} \sum_{i=1}^{n_d} \sum_{j=1}^D d_{ij} \log (\hat{d}_{ij}), \quad (4)$$

where n_d is the training data number, and D is the number of domains.

However, directly applying DANN to our specific problem of gait recognition does not work well. In fact, for WiFi signals, it is difficult to distinguish the features of different domains and the features of different gaits, which makes the classification of unknown target domains very challenging. To improve the classification performance of the target domains, we adopt two operations to optimize the model. Firstly, we concatenate \mathbf{Z}_i with the predicted label \hat{y}_i of the gait classification, and align the posterior distributions of the source and target domains:

$$P_i = Z_i \oplus \hat{y}_i, \quad (5)$$

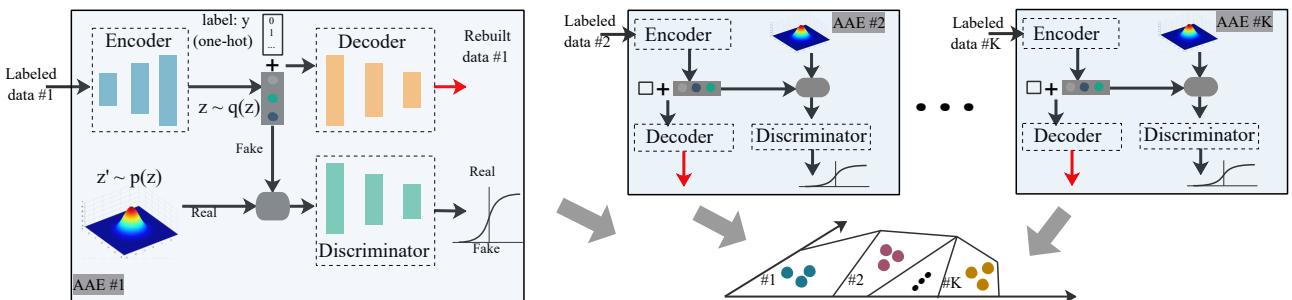


Fig. 9. We train AAE-based data augmentation with different distributions $N(0, k)$ for each user k to generate data similar to labeled data.

which together feed into the domain discriminator to predict the domain label \hat{d}_{ij}^o . We use cross-entropy as the loss function:

$$\mathcal{L}_o = -\frac{1}{n_d} \sum_{i=1}^{n_d} \sum_{j=1}^D d_{ij} \log (\hat{d}_{ij}^o). \quad (6)$$

Secondly, we add classification constraints to the target domains, and use the conditional entropy as the loss function for target domain classification:

$$\mathcal{L}_t = -\frac{1}{n_t} \sum_{i=1}^{n_t} \hat{d}_i^t \log (\hat{d}_i). \quad (7)$$

where n_t is the number of target domains data, \hat{d}_i^t is the predict label of target domain. In this way, the domain discriminator attempts to separate target domain data with the same domain label, thereby better obtaining domain-independent gait features.

Model training. The overall loss function of our model:

$$\mathcal{L}_{all} = \mathcal{L}_a + \alpha \mathcal{L}_o + \beta \mathcal{L}_d + \gamma \mathcal{L}_o + \lambda \mathcal{L}_t, \quad (8)$$

where α , β and γ are hyperparameters. The goal of model training is to minimize the loss $\mathcal{L}_a + \alpha \mathcal{L}_o$ of gait classification, while maximizing the loss $\beta \mathcal{L}_d + \gamma \mathcal{L}_o + \lambda \mathcal{L}_t$ of the domain discriminator. Note that the loss of the domain discrimination is inverted when backpropagated, while the gait classification is directly backpropagated [49]. We use all labeled and unlabeled data to train the model and iteratively update the parameters during the training process. In particular, we use a learning rate λ_p that changes as the iteration progresses so that the model learns the parameters better:

$$\lambda_p = \frac{\lambda_0}{(1 + \gamma \cdot p)^\mu} \quad (9)$$

where λ_0 is the initial learning rate, p is the ratio of the current to the total iterations, γ and μ are hyperparameters.

C. Data Augmentation

To better eliminate the influence of walking paths and behaviors on gait patterns, a feasible solution is to collect as much CSI as possible of different walking paths and behaviors. However, the time and labor costs of collecting and labeling data from different users are huge. In addition, although there are works to calculate the user's walking direction through two mutually perpendicular WiFi transceiver links [2], [18], [20], this requires strictly accurate prior position knowledge

of the transceiver and requires the user to walk normally and continuously in a specific area, so it cannot meet the needs of this paper. Fortunately, data augmentation schemes have been widely used recently, aiming to expand the training data set by generating more equivalent data from limited data [50]–[52]. Thus, we use the idea of data augmentation to generate more data on potential behaviors and paths.

Specifically, we design the AAE [28] as shown in Fig. 9, and use a separate AAE for each user's labeled gait data, that is, if K users need to be identified, we train K AAEs. AAE is a general method that can convert autoencoders into generative models. It combines adversarial ideas, and its typical network architecture consists of a standard *autoencoder* (AE) [53] and a *generative adversarial Network* (GAN) [54]. AAE aims to enable the decoder to generate realistic samples from any sampled data point by encouraging the encoder's output to completely fill the space of the prior distribution. GAN and AE are trained together in two stages, namely the reconstruction stage and the regularization stage. During the reconstruction stage, AE updates the encoder and decoder to minimize the reconstruction error of the input. In the regularization stage, GAN first updates its discriminator to separate real samples from generated samples, and then GAN updates its generator to confuse the discriminator. In this paper, we use a small amount of processed spectrogram data accompanying human behavior and different walking paths (*i.e.*, source domain data) for data augmentation, and we input these data S^k into the corresponding AAE $\#k$ for training. Firstly, the data S^k of the k th AAE is fed into the encoder of AAE $\#k$ to generate the latent vector $z \sim q(z)$, where $q(z)$ is the aggregate posterior distribution. z is sent to the decoder, and the vector \hat{S}^k is generated to reconstruct the data S^k . We define the reconstruction loss \mathcal{L}_B^k using *Mean Square Error* (MSE):

$$\mathcal{L}_B^k = \frac{1}{2n_g^k} \sum_{i=1}^{n_g^k} \left(S_i^k - \hat{S}_i^k \right)^2, \quad (10)$$

where n_g^k is the number of samples.

Secondly, we train the discriminator to normalize the rebuilt data. At this time, the encoder of AE becomes the generator of GAN, and its output S^k is sent to the discriminator together with the vector z' that obeys the prior distribution $p(z)$. Here, we choose the normal distribution $N(0, k)$ as the prior distribution $p(z)$. For the discriminator, the label l^k is 0 when S^k is used as input, and the label l^k is 1 when z' is used

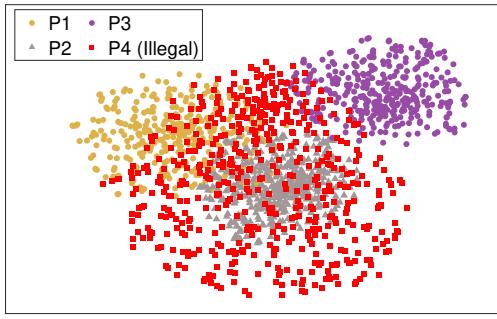


Fig. 10. Feature visualization of three legal users and one illegal user, where the features of the legal users are closely clustered, while those of the illegal user are randomly dispersed.

as input. Use cross entropy as the loss function \mathcal{L}_G^k of the discriminator:

$$\mathcal{L}_G^k = -\frac{1}{n_g^k} \sum_{i=1}^{n_g^k} \left(l_i^k \log(\hat{l}_i^k) + (1 - l_i^k) \log(1 - \hat{l}_i^k) \right). \quad (11)$$

However, we find that since we use separate AAE for each user for data augmentation, the rebuilt data samples are very likely to be located among different users, which in turn affects the classification accuracy of different users. In order to avoid this situation, as shown in Fig. 9, in addition to the prior distribution $p(z)$ of each AAE being different (the variance of the normal distribution is different), we also feed the user's label y (one-hot form) and the latent vector z into the decoder together to force the decoder to generate data that is highly relevant to users. Our goal for the augmented data is that the size of the rebuilt data for the k th user is 10 times larger than the size of the original labeled source domain data \mathbf{S}_L . Finally, we collect all rebuilt data \mathbf{S}_R as well as source domain data as labeled augmentation data $\mathbf{S}_A = \mathbf{S}_R \cup \mathbf{S}_L$.

For each user, an AAE is initialized with a small set of labeled gait data specific to that user, ensuring that the model captures unique gait characteristics. These AAEs are deployed in a parallelized, distributed framework, allowing for scalable training across multiple users without significantly increasing computational load. This setup leverages batch processing and asynchronous updates, enabling efficient use of computational resources and reducing overall training time as the number of users grows. During training, backpropagation is applied to minimize the reconstruction loss \mathcal{L}_B^k , ensuring accurate gait data reconstruction, with the goal of keeping this loss as low as possible. In the regularization stage, the discriminator's loss \mathcal{L}_G^k is targeted to be around 0.5, indicating a balance between real and generated samples.

D. Abnormal User Detection

Detecting abnormal users is a very necessary problem in actual gait recognition scenarios, because intruders may imitate the gait of legitimate users and try to attack. Therefore, we need to accurately identify abnormal users, even if there are no data samples of intruders in the training set. To analyze the extracted features of legitimate and illegal users, we first

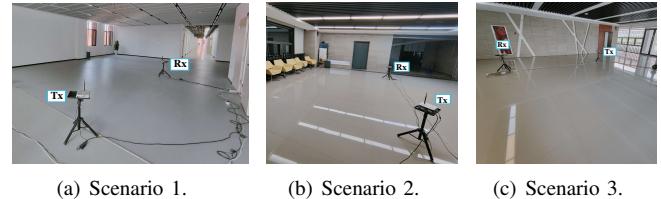


Fig. 11. Three different experimental scenarios.

train part of the data of three legitimate users (P1, P2, P3) to obtain a classification model. Then, we input the test data of the three legitimate users and the data of another illegal user (P4) into the model and visualize the features of the feature extraction module using t-SNE [55], and the results are shown in Fig. 10. We can see that the feature points of the three legitimate users are closely clustered, while the feature points of the illegal users (red points) are randomly distributed in the feature space. Therefore, we can implement abnormal user detection based on the idea of clustering. Specifically, during training, *freeGait* calculates the average neighbor distance of each legitimate category as the category density Θ_i :

$$\Theta_i = \frac{1}{K} \sum_{i=1}^K dis(Z_i) \quad (12)$$

where $dis(Z_i)$ is the i -th sample's feature. During testing, *freeGait* calculates the average distance $\theta_j = \frac{1}{K} \sum_{j=1}^K dis(Z_j)$ of the K nearest neighbors (KNN) of the test sample. If θ_j is greater than the category density Θ_τ of the most common category τ among its K neighbors, the test sample is considered to be from an illegal user, otherwise, the test sample is considered to be from a legitimate user:

$$L_j = \begin{cases} \tau, & \theta_j \leq \Theta_\tau \\ 0, & \theta_j > \Theta_\tau \end{cases} \quad (13)$$

where $L_j = 0$ indicates that sample j is from an illegal user.

V. PLATFORM IMPLEMENTATION

A. Hardware and Environments

We build the hardware platform based on two IPCs equipped with Intel 5300 NICs, one of them used as Tx with one antenna, another used as Rx with three antennas, and fixed on a tripod 0.5m above the ground to better detect human movement. As shown in Fig. 11, we deploy our experiments in three different environments, and the distance between Tx and Rx is 6m. Note that the environments and the positions of the WiFi transceiver remain unchanged during all tests. Although changes in the environment also cause changes in CSI, we do not analyze this in the paper. In addition, all data are preprocessed based on Matlab R2020b on a computer with an Intel-i5 2.7GHz CPU. The deep learning model training and result prediction are completed on a server with an NVIDIA 3090 graphics card based on Python 3.10 with CUDA 11.8 and Pytorch 2.0. We train our model offline with a total of 200 epochs. During training, the Adam optimizer is used with a learning rate of 0.001.

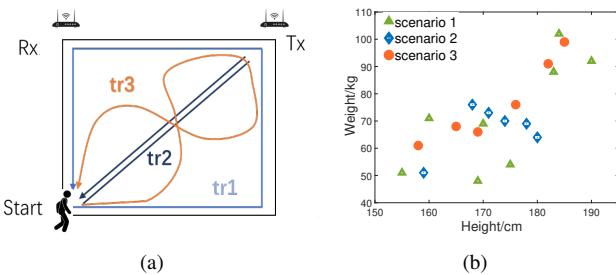


Fig. 12. Traces and volunteers: (a) Three different traces and (b) twenty volunteers with different heights and weights.

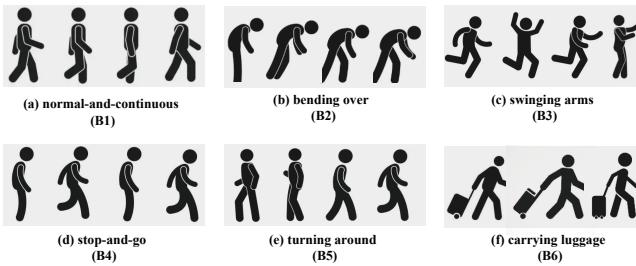


Fig. 13. We define six common human behaviors during walking as non-gait behaviors, including (a) normal-and-continuous (B1); (b) bending over (B2); (c) swinging arms (B3); (d) stop-and-go (B4); (e) turning around (B5); (f) carrying luggage (B6).

B. Data Collection

To collect CSI of different human behaviors and walking paths, we considered six human behaviors (*i.e.*, B1-B6) as shown in Fig. 13, and three paths (*i.e.*, tr1-tr3) as shown in Fig. 12(a). These behaviors were selected based on their potential to interact with and influence gait patterns, as each behavior may introduce different types of motion or perturbations to the walking process. These behaviors represent common, real-world actions that may occur during walking, making them well-suited for assessing the robustness of *freeGait* under diverse conditions. The six behaviors we defined are as follows: (a) Normal-and-continuous (B1): This behavior represents the baseline walking condition, without interruptions, allowing us to capture standard gait patterns. (b) Bending over (B2): Bending is a common behavior that can change body posture and affect weight distribution, which may affect gait patterns. (c) Swinging arms (B3): Arm movements affect the body's center of mass and affect walking rhythm, which may interfere with gait features. (d) Stop-and-go (B4): Sudden stops or changes in walking speed introduce perturbations that are particularly difficult for gait recognition systems to adapt to. (e) Turning around (B5): Turning involves a change in body orientation, which may cause noticeable changes in WiFi signals, affecting the accuracy of gait recognition. (f) Carrying luggage (B6): Carrying additional weight alters the body's natural movement patterns, making it a relevant behavior for testing how external factors affect gait recognition.

We recruit twenty volunteers of different heights and weights (*i.e.*, P1-P20) to record the CSI of different people in three scenarios as shown in Fig. 11. The allocation detail of scenarios and volunteers is as shown in Fig. 12(b). We install the CSI Tool on two IPCs and collect CSI us-

ing injection/monitoring mode with the sampling rate set to 1000Hz [30]. We ask each volunteer to walk and collect CSI as follows: (i) Each user walks along tr1 at a normal speed according to six different human behaviors, and each human behavior is repeated three times. (ii) Each user walks along tr2 at a normal speed according to six different human behaviors, and each human behavior is repeated three times. (iii) Each user walks along tr3 according to human behavior 1 (*i.e.*, B1) at three different speeds (*i.e.*, slow (S1), normal (S2), fast (S3)) for one minute at each speed.

Then, we follow the method in Section IV-A to perform data preprocessing and obtain the spectrograms. Specifically, there are $\sum_{i=1}^{N_b} N_i$ sets of spectrogram data for each user (N_i is the number of sliced data in each case, N_b is the total number of cases, *i.e.*, the number of domains), and the size of each group of data is $3 \times 80 \times 458$ (antenna \times frequency \times time). We divide the data into $6 + 6 + 3 = 15$ domains based on walking paths and human behaviors, and randomly divide the data in each domain into 50% training data and 50% testing data. Furthermore, for the training data, the source domain data has user labels, while the target domain data does not have user labels. Additionally, all tags are encoded using one-hot.

C. Baseline Methods

- **Basic-CNN** is used as the baseline, *i.e.*, without data augmentation and domain adaptation, labeled and unlabeled data are directly input into CNN and LSTM-based feature extractors and gait classifiers to directly predict users' labels. Some previous work is based on this scheme [2].
- **AAE-only** implements only *freeGait*'s AAE and connects them directly to the feature extractor and gait classifier without domain adaptation.
- **Domain-adaptation-only** implements only the domain adaptation network of *freeGait*, *i.e.*, directly feeds labeled and unlabeled data into the feature extractor, gait classifier, and domain discriminator without AAEs.
- **WiFi-ID** [1] employs manually designed time-domain and frequency-domain features and uses the ReliefF algorithm to select high-weight features for classification based on Sparse Approximation Classification (SAC).
- **WifiU** [17] applies a series of spectrogram enhancement techniques and manual feature extraction to serve as fingerprints for user gait patterns, and trains a LibSVM classifier.

Notably, we avoid comparing our system with gait recognition methods based on multi-transceiver setups (*e.g.*, WiDGR [18], Wi-PIGR [2], MetaGANFi [19]) because our data is collected using a single transceiver pair. MetaGANFi, in particular, is designed for environments with multiple transceivers and aims to address device position changes, which is fundamentally different from our single-transceiver approach. Therefore, direct comparison with these systems is not fair or applicable to our work.

VI. EVALUATION

A. Basic Performance of *freeGait*

To evaluate the basic performance of *freeGait*, we first classify the gait data of three volunteers in Scenario 1, as

True Class	P1	23.3333	0
P1	76.6667	23.3333	0
P2	33.0508	66.9492	0
P3	98.088	1.91205	0
Predict Class	P1	P2	P3

(a) Basic-CNN.

True Class	P1	15.5556	35.5556
P1	48.8889	15.5556	35.5556
P2	27.1186	48.3051	24.5763
P3	0.573614	0.573614	98.8528
Predict Class	P1	P2	P3

(b) AAE-only.

True Class	P1	15.5556	1.11111
P1	83.3333	15.5556	1.11111
P2	56.9915	41.1017	1.90678
P3	7.26577	1.14723	91.587
Predict Class	P1	P2	P3

(c) Domain-adaptation-only.

True Class	P1	4.72222	3.88889
P1	91.3889	4.72222	3.88889
P2	6.69492	89.4915	3.81356
P3	2.10325	1.91205	95.9847
Predict Class	P1	P2	P3

(d) *freeGait*.

Fig. 14. The user prediction confusion matrices obtained after the predicted data of three users pass through the basic CNN, AAE-only, domain adaptation-only, and *freeGait* respectively.

TABLE I

DATASET PARTITIONING FOR BASIC PERFORMANCE, THE IMPACT OF DIFFERENT BEHAVIORS, TRACES, SPEED, AMOUNT OF AUGMENTED DATA, SAMPLING RATE, AND LOSS FUNCTION.

Scenario	Scenario 1
People	P1, P2, P3
Source Domains	[B1/B4/B6_tr1, B5_tr2, S1/S3_tr3] (50%)
Target Domains	Other 9 Combinations (50%)
Test Set	All 15 Combinations (50%)

TABLE II

CLASSIFICATION RESULTS OF THE SIX METHODS.

Method	TPR	FPR
WiFi-ID	56.93%	21.94%
WifiU	61.86%	19.07%
Basic-CNN	47.87%	26.06%
AAE-only	65.35%	17.33%
Domain-adaptation-only	72.01%	14.00%
<i>freeGait</i>	92.29%	3.86%

shown in Tab. I. We select 6 of the 15 combinations of human behaviors, traces, and speeds of these three volunteers in the training data as source domains data (B1_tr1, B4_tr1, B6_tr1, B5_tr2, S1_tr3, S3_tr3), *i.e.*, with user labels, and the remaining training data is used as target domains data, *i.e.*, without user labels, and all training data contains domain labels. During the training process, both Domain-adaptation-only and *freeGait* can train their respective models using data with user labels and data without user labels, but Basic-CNN and AAE-only just use data with user labels to train their models. We then compare *freeGait* with the above baseline methods.

We use True Positive Rate (TPR) and False Positive Rate (FPR) to evaluate classification performance, where $TPR = \frac{TP}{TP+FN}$, and $FPR = \frac{FP}{TN+FP}$, TP, TN, FP, and FN represent the number of true positives, true negatives, false positives, and false negatives. Tab. II shows the classification results of six methods. As can be seen, both WiFi-ID [1] and WifiU [17] exhibit significant accuracy degradation when gait features are confused due to interference from non-gait behaviors. The accuracy of the Base-CNN is extremely low because it only uses less labeled data for training. AAE-only and Domain-adaption-only improve some accuracy through data augmentation and domain adaptation respectively, but there is still a lot of room for improvement. Compared with these techniques, *freeGait* can improve accuracy by at least 20% and in the worst case 45%.

To more intuitively show the comparison results of *freeGait* with other methods, we also provide the gait classification

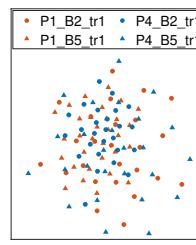
True Class	P1	20	0
P1	80	20	0
P2	62.9237	36.8644	0.211864
P3	41.3002	4.78011	53.9197
Predict Class	P1	P2	P3

(a) WiFi-ID.

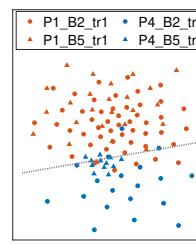
True Class	P1	68.3333	4.16667
P1	27.5	68.3333	4.16667
P2	22.2458	72.0339	5.72034
P3	5.92734	8.03059	86.0421
Predict Class	P1	P2	P3

(b) WiFiU.

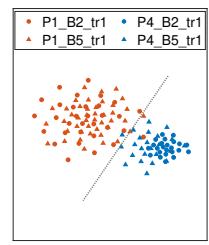
Fig. 15. The user prediction confusion matrices obtained after the predicted data of three users pass through WiFi-ID and WiFiU.



(a) Basic-CNN.



(b) DANN-only.



(c) *freeGait*.

Fig. 16. Visualization results of gait features after going through Basic-CNN, DANN-only and *freeGait*.

confusion matrix, and the results are shown in Fig. 14 and Fig. 15. From the results, compared to other schemes that cannot accurately distinguish multiple users, *freeGait* can achieve high classification accuracy for all users.

B. Visualization of Gait Feature

Our model aims to learn representations of gait features that are independent of discontinuous human behaviors and walking paths. To verify that the model has learned the representation, we use t-SNE [55] to reduce the dimensionality and display it in 2D space. Specifically, we select the data of two human behaviors of two users from the data in the target domain that does not carry user tags, *i.e.*, four domain-user pairs. Then, we randomly select several samples for each domain-user pair and draw the learned representations of these samples through Basic-CNN, DANN-only (*i.e.*, Domain-adaptation-only) and *freeGait* respectively. The results are shown in Fig. 16, where different colors (orange and blue) represent different users, and different shapes (circles and triangles) represent different human behaviors. When Basic-CNN is used to extract gait features, the gait features of two users overlap. Although DANN-only can distinguish the gait features of two users, these features are scattered and not concentrated into two clusters. In addition, the gait features extracted by *freeGait* are concentrated in two clusters, while

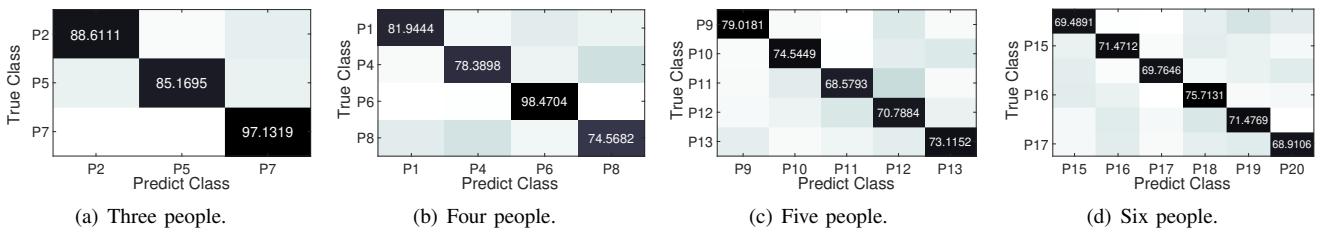


Fig. 17. Classification results for different people and different numbers of people.

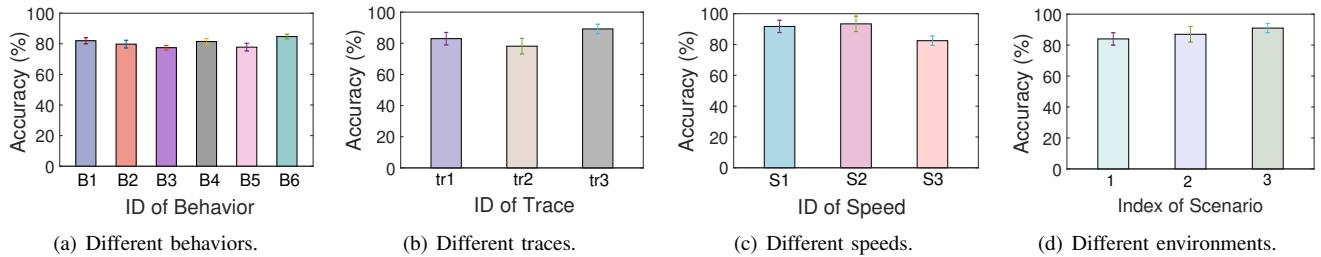


Fig. 18. Classification accuracy for different behaviors, traces, speeds, and environments.

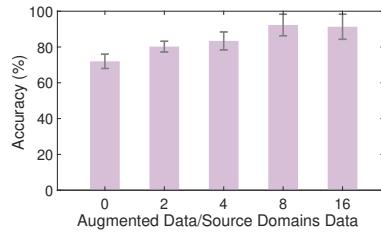


Fig. 19. Impact of augmented data amount.

there is almost no overlap between different users. This proves that the proposed model learns the target features.

C. Impact of People's Number

The increase in the number of users and different users affects the classification accuracy of gait features. To further examine the performance of *freeGait*, we test *freeGait* on different users and a larger number of users. Specifically, we divide 20 volunteers into four different group sizes of 3, 4, 5, and 6 people. The confusion matrix for classification is shown in Fig. 17. It can be seen that as the number of people increases, *freeGait* can still achieve 70% classification accuracy for each user, but the overall classification accuracy decreases slightly, because the features of domains are also gradually increasing. We plan to optimize our model in future work to be able to predict more diverse users.

D. Impact of Different Behaviors, Traces and Speeds

Different human behaviors, walking paths, and walking speeds all affect gait patterns. In order to evaluate the robustness of *freeGait* to these influencing factors, we divide the users' gait data into 15 groups according to domain categories, and predict users' IDs respectively (the source domain data for training is the same as Section VI-A), the accuracy is shown in Fig. 18. The results show that for the six (behavior, speed, trace) including the source domains, the users' gait classification accuracy is above 81%, while for the nine (behavior, speed, trace) that are all in the target domains,

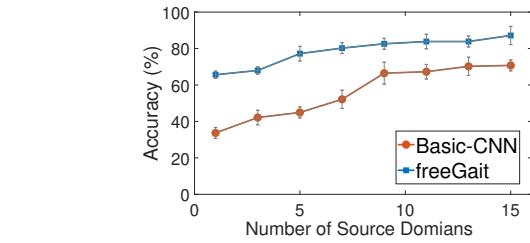


Fig. 20. Impact of source domains' number.

the users' gait classification accuracy still higher than 72%, although it has declined. Therefore, *freeGait* remains robust to different human behaviors, walking paths and walking speeds.

E. Impact of Different Environments

Different environments affect the collected CSI data. In order to evaluate the robustness of *freeGait* to different environments, we divide the data into three groups according to the three scenarios as shown in Fig. 11, and predict users' IDs respectively (the source domain data for training is the same as Section VI-A). The results in Fig. 18(d) show that for the three scenarios, the users' gait classification accuracy is above 83%, thus *freeGait* remains robust to different environments. It is worth noting that we evaluate *freeGait* in three scenarios separately without predicting users' IDs across environments, which is beyond the scope of this paper.

F. Impact of Augmented Data Amount

We already know that augmented data can improve the performance of *freeGait*. To analyze in detail the effect of the amount of augmented data on the users' gait classification effect, we select three users in Scenario 1 for verification. Specifically, we use pre-trained AAE-based data augmentation models to generate augmented data that are 0 times, 2 times, 4 times, 8 times, and 16 times the amount of source domain data, respectively. Then, we use these data to train *freeGait* respectively, and the accuracy of users' gait classification is shown in Fig. 19. The results show that the classification

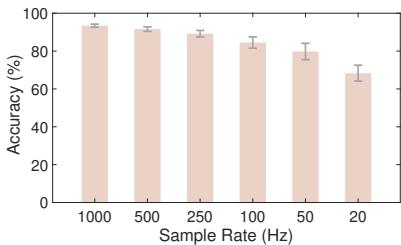


Fig. 21. Classification accuracy for different sampling rate.

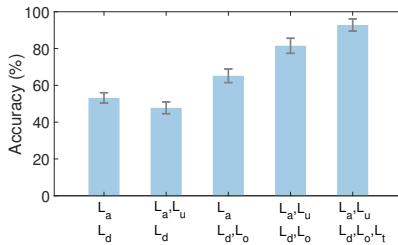


Fig. 22. Classification accuracy for different loss functions.

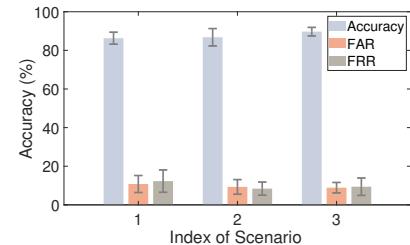


Fig. 23. Detection accuracy of abnormal user for different scenarios.

accuracy increases as the amount of augmented data increases. However, when there is too much augmented data (16 times), the classification accuracy decreases slightly. This may be caused by excessive enhancement that makes the sample distribution uneven.

G. Impact of Source Domains' Number

Intuitively, the amount of source domains used for training affects the accuracy of the model. To analyze the impact in detail, we pre-train the deep learning model of *freeGait* and predict users' IDs with the number of source domains ranging from 1 to 15. The result is shown in Fig. 20. It can be seen that the users' gait classification accuracy increases as the number of source domains increases, and the classification accuracy of *freeGait* is always higher than Basic-CNN with the same number of source domains. A limited number of source domains may result in overfitting to specific features of those domains or underfitting, where the model fails to learn sufficient patterns to generalize well to new data. This can also lead to instability in the training process, as the model may not have enough variability in the training data to effectively capture the underlying patterns of human gait. In practical scenarios where data is scarce, small sample sizes can also make the model sensitive to noise or biases inherent in the data. In real-world applications, users can choose an appropriate number of source domains by considering the balance between accuracy and training data collection cost. When data is limited, techniques such as data augmentation and transfer learning provide practical solutions to enhance model performance and stability.

H. Impact of Sampling Rate

In real-world scenarios, CSI can be estimated during active data communication. When data communication is inactive, the sampling rate of CSI significantly decreases. To analyze the impact of different sampling rates, we randomly sampled data from the original 1000Hz to simulate cases with sampling rates of 500Hz, 250Hz, 100Hz, 50Hz, and 20Hz. We then train our models using the data with these lower sampling rates and use them to predict user identities. The results, as shown in Fig. 21, indicate that prediction accuracy does not decline significantly from 1000Hz to 250Hz. However, there is a sharp drop in accuracy below 100Hz. This decline occurs because the frequency of human gait ranges from 10Hz to 70Hz. When the sampling rate is too low, a substantial amount of valuable

frequency information is lost, leading to a significant reduction in prediction accuracy.

I. Impact of Loss Function

In addition, we also verify the impact of the loss function in our proposed model design, namely L_u , L_o , and L_t . Specifically, we train the model for the following five cases in the loss function: (L_a, L_d) , (L_a, L_u, L_d) , (L_a, L_d, L_o) , (L_a, L_u, L_d, L_o) and $(L_a, L_u, L_d, L_o, L_t)$, and then use the trained model to predict users. As shown in Fig. 22, the prediction accuracy of users without the three constraints is much lower than that with the three constraints. This shows that it is not feasible to directly use the original DANN network [49] to train our data, so *freeGait* is necessary.

J. Accuracy of Abnormal User Detection

Finally, we evaluate the performance of abnormal user detection. First, for three different scenarios, we randomly select four users to construct a legal user dataset, and one user as an illegal user. Then, we train the model using the training data of legal users, and predict the test data of legal users and the data of illegal users. We use accuracy, *False Alarm Rate* (FAR), and *False Rejection Rate* (FRR) to evaluate the results, where FAR is defined as the ratio of illegal users classified as legal users, and FRR is defined as the ratio of legal users classified as illegal users. The results are shown in Fig. 23. In all three scenarios, our abnormal user detection module can achieve an accuracy of more than 85%, while FAR and FRR are both less than 12%.

VII. DISCUSSIONS AND FUTURE WORK

Diverse real-world non-gait behaviors. In the real world, human non-gait behaviors are far more diverse and complex than the six behaviors we have defined. There are many similar or unpredictable actions, such as falling, lying down, and jumping, that extend beyond our initial scope. Verifying these broader behaviors is crucial for the advancement of *freeGait*. To address this, we plan to adopt an iterative model training approach in future work, enabling us to continuously update the non-gait behavior library and enhance *freeGait*'s adaptability to a wider range of previously unseen human behaviors. Furthermore, we aim to conduct a detailed analysis of various human behaviors and develop metrics to evaluate non-gait behaviors, providing deeper insights into how different non-gait behaviors impact gait recognition systems.

Specifically, we could define two dimensions: "amplitude" and "complexity" to distinguish between large movements (such as jumping or rapid gestures) and small, subtle motions (like minor hand movements or slight posture shifts). Additionally, we envision designing a set of comprehensive metrics, such as "behavioral interference" and "gait recognition error rate" to quantify how various non-gait behaviors influence the performance of the gait recognition system. These metrics would allow us to gain deeper insights into the specific ways different behaviors disrupt gait features, thereby providing valuable guidance for future model improvements and non-gait behavior classification. This approach will also support the continuous expansion of our non-gait behavior library, enhancing the system's ability to adapt to previously unseen behaviors.

Diverse scenarios and walking paths. This paper focuses on addressing the core challenge of mitigating non-gait behaviors in WiFi-based gait recognition systems. To validate the effectiveness of the proposed method, we conducted experiments in controlled environments with three representative walking paths. These paths were chosen because they reflect common patterns in practical applications, such as hallways, offices, and indoor monitoring scenarios, and provide a solid foundation for evaluating the system. However, we acknowledge that these controlled scenarios may not fully capture the diversity and complexity of real-world environments. In future work, we plan to expand our evaluation to include cluttered spaces, dynamic settings with moving obstacles, and varied environmental conditions to test the system's robustness. Additionally, we aim to refine the system to handle unconstrained walking paths, allowing for more natural movement patterns and improving its applicability to realistic scenarios.

Cross-environment gait recognition. In this paper, we do not address the migration of the same user across different environments, as our focus is on examining the impact of human behaviors on gait recognition systems. However, in real-world scenarios, cross-environment gait recognition is indeed a critical research challenge due to environmental complexity. We plan to explore this area in future work. Specifically, we aim to accurately distinguish between different or even similar environments, treating the environment as a domain, so that cross-domain techniques can be applied to achieve cross-environment gait recognition. Additionally, we intend to develop a WiFi-based gait recognition system that simultaneously considers factors such as human behaviors, trajectories, and environmental conditions, facilitating the real-world deployment of such solutions.

Multiple users identification. The solutions proposed so far are primarily focused on the gait recognition of a single individual. However, given the prevalence of multi-person scenarios in real-world applications, recognizing the gaits of multiple users simultaneously is of great importance. Due to the nature of WiFi signals, traditional model-driven approaches struggle to effectively separate the gait data of different users, particularly when two users have similar walking frequencies. Fortunately, recent studies have demonstrated that data-driven approaches perform well in estimating WiFi sensing data with multiple frequency aliasing [56], [57]. Inspired by these

advancements, we can incorporate simulated data generation based on electromagnetic field models or diffusion models into the data augmentation module, thereby enabling the recognition of multiple users simultaneously.

Data of different WiFi NICs. In this paper, our gait dataset is constructed using an Intel 5300 NIC, leveraging the *Linux 802.11n CSI Tool* [30]. However, there are other tools available that can collect CSI from various WiFi NICs for diverse sensing tasks, such as the *Atheros CSI Tool* [31], *Nexmon CSI Extractor* [32], and *PicoScenes* [58]. Additionally, some researchers are dedicated to building comprehensive WiFi sensing datasets using different WiFi devices [59]–[61]. To further advance our system for practical applications, it is essential to collect CSI data and evaluate our model across different WiFi NICs and heterogeneous devices. In future work, we plan to continue gathering gait data using various WiFi NICs.

VIII. RELATED WORK

A. WiFi-based Gait Recognition Systems

Compared to vision-based [4], acoustic signal-based [7], and wearable device-based [9] gait recognition technologies, WiFi has garnered significant attention recently due to its ubiquity, non-contact nature, and non-intrusive privacy advantages. Numerous systems have been developed that accurately recognize human gait using commercial WiFi devices. However, most of these systems require subjects to walk along specific paths [1], [17], [19], [62]–[64], and some combine daily activity features to enhance accuracy [65], [66]. These approaches either impose strict constraints on user behavior or necessitate the collection of large amounts of data. Recent studies have explored gait recognition methods that are independent of walking direction and path, but these still require data from multiple WiFi transceivers and necessitate that the subject walks normally and continuously within a specific area [2], [18], [67]. In contrast, *freeGait* is designed to handle a variety of discontinuous behaviors that may accompany a user's walking. It accurately identifies the user's gait by collecting only a small amount of data from a pair of randomly placed WiFi transceivers, while also relaxing the constraints on the user's walking requirements.

B. Domain adaptation for WiFi sensing

WiFi signals are highly sensitive to their surrounding environment. Variations in the environment, different individuals, and diverse actions can all alter the reflected signal information, which in turn affects the performance of WiFi-based sensing systems. Collecting sufficient data to address these variations is both costly and impractical. To mitigate this issue, domain adaptive learning enhances the learning performance in data-scarce target domains by minimizing the distribution differences between the source and target domains [49]. As a result, this approach has increasingly been applied in WiFi sensing systems, including applications such as fingerprint localization [22], [68], object recognition [27], [69], activity recognition [23], [70], [71], gesture recognition [24]–[26], and human identification [19]. However, existing domain

adaptation works in WiFi sensing mainly target challenges related to changes in environment or device setup. While these approaches achieve robustness in their respective tasks, they do not address the unique challenge of behavioral diversity in gait recognition. In contrast, our work introduces a novel approach to align the posterior distribution to mitigate classification ambiguity caused by non-gait behaviors, significantly interfering with gait feature extraction. In addition, in this paper, we integrate the concept of domain adaptation with data augmentation to enable our system to adapt to the various non-periodic or discontinuous behaviors that may accompany a user's walking. This approach allows the WiFi-based gait recognition system to maintain accuracy even when confronted with non-gait behaviors.

IX. CONCLUSION

In this paper, we present *freeGait*, a WiFi-based gait recognition system designed to remain robust against non-periodic and discontinuous user behaviors as well as varying walking paths. *freeGait* introduces a novel approach to eliminate CSI noise, enabling the extraction of fine-grained spectrograms. It employs a deep learning framework that integrates data augmentation and domain adaptation to address inconsistencies in gait patterns resulting from diverse user behaviors and paths. To ensure domain adaptation effectively addresses our specific challenges, we align the posterior distributions between the source and target domains and constrain the conditional distribution of the target domains to optimize the model. For data augmentation, we tackle the issue of aliasing in reconstructed samples by utilizing supervised learning, which guides the decoder to generate data highly relevant to the user. By leveraging a large volume of unlabeled data from a pair of WiFi transceivers, *freeGait* enables the model to learn user gait features that are independent of human behaviors and paths, relying on only a small amount of labeled data. Our experiments, conducted with 20 volunteers across three real-world scenarios, demonstrate that *freeGait* outperforms existing methods in handling complex unlabeled human behaviors and walking paths. This capability to mitigate the impact of non-gait behaviors significantly enhances the potential for practical deployment of WiFi-based gait recognition systems.

ACKNOWLEDGMENTS

The research is partially supported by the "Pioneer" and "Leading Goose" R&D Program of Zhejiang (Grant No. 2023C01029), NSFC with No. 62072424, U20A20181, Key Research Program of Frontier Sciences, CAS. No. ZDBS-LY-JSC001, Anhui Provincial Natural Science Foundation with No. 2308085MF221, Hefei Municipal Natural Science Foundation with No. 2022016, the Fundamental Research Funds for the Central Universities with No. WK3500000008.

REFERENCES

- [1] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "Wifi-id: Human identification using wifi signal," in *2016 International Conference on Distributed Computing in Sensor Systems (DCOSS)*. IEEE, 2016, pp. 75–82.
- [2] L. Zhang, C. Wang, and D. Zhang, "Wi-pigr: Path independent gait recognition with commodity wi-fi," *IEEE Transactions on Mobile Computing*, vol. 21, no. 9, pp. 3414–3427, 2021.
- [3] L. Console, I. Torre, I. Lombardi, S. Gioria, and V. Surano, "Personalized and adaptive services on board a car: an application for tourist information," *Journal of Intelligent Information Systems*, vol. 21, pp. 249–284, 2003.
- [4] H. Chao, Y. He, J. Zhang, and J. Feng, "Gaitset: Regarding gait as a set for cross-view gait recognition," 2018.
- [5] C. Wang, J. Zhang, L. Wang, J. Pu, and X. Yuan, "Human identification using temporal information preserving gait template," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2164–2176, 2011.
- [6] Y. Wang, Y. Chen, M. Z. A. Bhuiyan, Y. Han, S. Zhao, and J. Li, "Gait-based human identification using acoustic sensor and deep neural network," *Future Generation Computer Systems*, vol. 86, pp. 1228–1237, 2018.
- [7] W. Xu, Z. Yu, Z. Wang, B. Guo, and Q. Han, "Acousticid: gait-based human identification using acoustic signal," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, pp. 1–25, 2019.
- [8] D. Gafurov, "A survey of biometric gait recognition: Approaches, security and challenges," in *Annual Norwegian computer science conference*. Annual Norwegian Computer Science Conference Norway, 2007, pp. 19–21.
- [9] M. Muaz and R. Mayrhofer, "Smartphone-based gait recognition: From authentication to imitation," *IEEE Transactions on Mobile Computing*, vol. 16, no. 11, pp. 3209–3221, 2017.
- [10] X. Yang, J. Liu, Y. Chen, X. Guo, and Y. Xie, "Mu-id: Multi-user identification through gaits using millimeter wave radios," in *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*. IEEE, 2020, pp. 2589–2598.
- [11] D. Wang, X. Zhang, K. Wang, L. Wang, X. Fan, and Y. Zhang, "Rdgait: A mmwave based gait user recognition system for complex indoor environments using single-chip radar," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 8, no. 3, pp. 1–31, 2024.
- [12] P. Airey and J. Verran, "A method for monitoring substratum hygiene using a complex soil: The human fingerprint," *Fouling, cleaning and disinfection in food processing*, 2006.
- [13] Y. E. Du, "Review of iris recognition: cameras, systems, and their applications," *Sensor review*, vol. 26, no. 1, pp. 66–69, 2006.
- [14] Z. Xue, D. Ming, W. Song, B. Wan, and S. Jin, "Infrared gait recognition based on wavelet transform and support vector machine," *Pattern recognition*, vol. 43, no. 8, pp. 2904–2910, 2010.
- [15] D. Kim, S. Lee, and J. Paik, "Active shape model-based gait recognition using infrared images," in *Signal Processing, Image Processing and Pattern Recognition: International Conference, SIP 2009, Held as Part of the Future Generation Information Technology Conference, FGIT 2009, Jeju Island, Korea, December 10–12, 2009. Proceedings*. Springer, 2009, pp. 275–281.
- [16] D. Ming, Z. Xue, L. Meng, B. Wan, Y. Hu, and K. D. Luk, "Identification of humans using infrared gait recognition," in *2009 IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurements Systems*. IEEE, 2009, pp. 319–322.
- [17] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using wifi signals," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 363–373.
- [18] L. Zhang, C. Wang, M. Ma, and D. Zhang, "Widigr: Direction-independent gait recognition system using commercial wi-fi devices," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1178–1191, 2019.
- [19] J. Zhang, Z. Chen, C. Luo, B. Wei, S. S. Kanhere, and J. Li, "Metaganfi: Cross-domain unseen individual identification using wifi signals," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 6, no. 3, pp. 1–21, 2022.
- [20] D. Wu, D. Zhang, C. Xu, Y. Wang, and H. Wang, "Widir: Walking direction estimation using wireless signals," in *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*, 2016, pp. 351–362.
- [21] Q. Bu, X. Ming, J. Hu, T. Zhang, J. Feng, and J. Zhang, "Transfersense: towards environment independent and one-shot wifi sensing," *Personal and Ubiquitous Computing*, pp. 1–19, 2021.
- [22] H. Li, X. Chen, J. Wang, D. Wu, and X. Liu, "Dafi: Wifi-based device-free indoor localization via domain adaptation," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 4, pp. 1–21, 2021.

- [23] W. Jiang, C. Miao, F. Ma, S. Yao, Y. Wang, Y. Yuan, H. Xue, C. Song, X. Ma, D. Koutsoukos *et al.*, "Towards environment independent device free human activity recognition," in *Proceedings of the 24th annual international conference on mobile computing and networking*, 2018, pp. 289–304.
- [24] R. Xiao, J. Liu, J. Han, and K. Ren, "Onefi: One-shot recognition for unseen gesture via cots wifi," in *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, 2021, pp. 206–219.
- [25] D. Wang, J. Yang, W. Cui, L. Xie, and S. Sun, "Airfi: empowering wifi-based passive human gesture recognition to unseen environment via domain generalization," *IEEE Transactions on Mobile Computing*, vol. 23, no. 2, pp. 1156–1168, 2022.
- [26] H. Zou, J. Yang, Y. Zhou, L. Xie, and C. J. Spanos, "Robust wifi-enabled device-free gesture recognition via unsupervised adversarial domain adaptation," in *2018 27th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 2018, pp. 1–8.
- [27] E. Soltanaghaei, R. A. Sharma, Z. Wang, A. Chittilappilly, A. Luong, E. Giler, K. Hall, S. Elias, and A. Rowe, "Robust and practical wifi human sensing using on-device learning with a domain adaptive model," in *Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, 2020, pp. 150–159.
- [28] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, "Adversarial autoencoders," *arXiv preprint arXiv:1511.05644*, 2015.
- [29] D. Yan, Y. Yan, P. Yang, W.-Z. Song, X.-Y. Li, and P. Liu, "Real-time identification of rogue wifi connections in the wild," *IEEE Internet of Things Journal*, vol. 10, no. 7, pp. 6042–6058, 2022.
- [30] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11 n traces with channel state information," *ACM SIGCOMM computer communication review*, vol. 41, no. 1, pp. 53–53, 2011.
- [31] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity wifi," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '15. New York, NY, USA: ACM, 2015, p. 53–64. [Online]. Available: <http://doi.acm.org/10.1145/2789168.2790124>
- [32] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free your csi: A channel state information extraction platform for modern wi-fi chipsets," in *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*, ser. WiNTECH '19, 2019, p. 21–28. [Online]. Available: <https://doi.org/10.1145/3349623.3355477>
- [33] F. Han, C. Wan, P. Yang, H. Zhang, Y. Yan, and X. Cui, "Ace: Accurate and automatic csi error calibration for wireless localization system," in *2020 6th International Conference on Big Data Computing and Communications (BIGCOM)*. IEEE, 2020, pp. 15–23.
- [34] X. Niu, S. Li, Y. Zhang, Z. Liu, D. Wu, R. C. Shah, C. Tanriover, H. Lu, and D. Zhang, "Wimonitor: Continuous long-term human vitality monitoring using commodity wi-fi devices," *Sensors*, vol. 21, no. 3, p. 751, 2021.
- [35] C. Chen, Y. Han, Y. Chen, H.-Q. Lai, F. Zhang, B. Wang, and K. R. Liu, "Tr-breath: Time-reversal breathing rate estimation and detection," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 3, pp. 489–501, 2017.
- [36] F. Zhang, C. Wu, B. Wang, H.-Q. Lai, Y. Han, and K. R. Liu, "Widetect: Robust motion detection with a statistical electromagnetic model," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, pp. 1–24, 2019.
- [37] Y. Ma, G. Zhou, and S. Wang, "Wifi sensing with channel state information: A survey," *ACM Computing Surveys (CSUR)*, vol. 52, no. 3, pp. 1–36, 2019.
- [38] L. I. Smith, "A tutorial on principal components analysis," Cornell University, USA, Tech. Rep., February 26 2002. [Online]. Available: http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf
- [39] Y. Zhang, F. Han, P. Yang, Y. Feng, Y. Yan, and R. Guan, "Wicyclops: Room-scale wifi sensing system for respiration detection based on single-antenna," *ACM Transactions on Sensor Networks*, 2023.
- [40] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang, "Farsense: Pushing the range limit of wifi-based respiration sensing with csi ratio of two antennas," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, pp. 1–26, 2019.
- [41] S. Li, Z. Liu, Y. Zhang, Q. Lv, X. Niu, L. Wang, and D. Zhang, "Wiborder: Precise wi-fi based boundary sensing via through-wall discrimination," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 3, pp. 1–30, 2020.
- [42] J. Liu, Y. Zeng, T. Gu, L. Wang, and D. Zhang, "Wiphone: Smartphone-based respiration monitoring using ambient reflected wifi signals," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 1, pp. 1–19, 2021.
- [43] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, ser. KDD'96. AAAI Press, 1996, p. 226–231.
- [44] N. Y. Hammerla, S. Halloran, and T. Plötz, "Deep, convolutional, and recurrent models for human activity recognition using wearables," *arXiv preprint arXiv:1604.08880*, 2016.
- [45] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," *Neural computation*, vol. 12, no. 10, pp. 2451–2471, 2000.
- [46] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.
- [47] Y. Grandvalet and Y. Bengio, "Semi-supervised learning by entropy minimization," *Advances in neural information processing systems*, vol. 17, 2004.
- [48] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, 2009.
- [49] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," *The journal of machine learning research*, vol. 17, no. 1, pp. 2096–2030, 2016.
- [50] A. Antoniou, A. Storkey, and H. Edwards, "Data augmentation generative adversarial networks," *arXiv preprint arXiv:1711.04340*, 2017.
- [51] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of big data*, vol. 6, no. 1, pp. 1–48, 2019.
- [52] C. Shorten, T. M. Khoshgoftaar, and B. Furht, "Text data augmentation for deep learning," *Journal of big Data*, vol. 8, pp. 1–34, 2021.
- [53] A. Ng *et al.*, "Sparse autoencoder," *CS294A Lecture notes*, vol. 72, no. 2011, pp. 1–19, 2011.
- [54] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.
- [55] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. 11, 2008.
- [56] Z. Yang, Y. Zhang, K. Qian, and C. Wu, "SLNet: A spectrogram learning neural network for deep wireless sensing," in *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*. Boston, MA: USENIX Association, Apr. 2023, pp. 1221–1236. [Online]. Available: <https://www.usenix.org/conference/nsdi23/presentation/yang-zheng>
- [57] G. Chi, Z. Yang, C. Wu, J. Xu, Y. Gao, Y. Liu, and T. X. Han, "Rf-diffusion: Radio signal generation via time-frequency diffusion," in *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*, 2024, pp. 77–92.
- [58] Z. Jiang, T. H. Luan, X. Ren, D. Lv, H. Hao, J. Wang, K. Zhao, W. Xi, Y. Xu, and R. Li, "Eliminating the barriers: Demystifying wi-fi baseband design and introducing the picoscenes wi-fi sensing platform," *IEEE Internet of Things Journal*, vol. 9, no. 6, pp. 4476–4496, 2021.
- [59] F. Wang, Y. Lv, M. Zhu, H. Ding, and J. Han, "Xrf55: A radio frequency dataset for human indoor action analysis," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 8, no. 1, pp. 1–34, 2024.
- [60] Y. Zheng, Y. Zhang, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, "Zero-effort cross-domain gesture recognition with wi-fi," in *Proceedings of the 17th annual international conference on mobile systems, applications, and services*, 2019, pp. 313–325.
- [61] J. Yang, H. Huang, Y. Zhou, X. Chen, Y. Xu, S. Yuan, H. Zou, C. X. Lu, and L. Xie, "Mm-fi: Multi-modal non-intrusive 4d human dataset for versatile wireless sensing," *Advances in Neural Information Processing Systems*, vol. 36, 2024.
- [62] J. Zhang, Z. Tang, M. Li, D. Fang, P. Nurmi, and Z. Wang, "Crosssense: Towards cross-site and large-scale wifi sensing," in *Proceedings of the 24th annual international conference on mobile computing and networking*, 2018, pp. 305–320.
- [63] Y. Zeng, P. H. Pathak, and P. Mohapatra, "Wiwho: Wifi-based person identification in smart spaces," in *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 2016, pp. 1–12.
- [64] J. Zhang, B. Wei, F. Wu, L. Dong, W. Hu, S. S. Kanhere, C. Luo, S. Yu, and J. Cheng, "Gate-id: Wifi-based human identification irrespective of

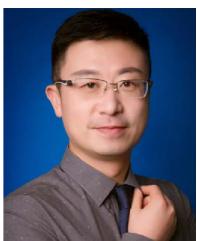
- walking directions in smart home," *IEEE Internet of Things Journal*, vol. 8, no. 9, pp. 7610–7624, 2020.
- [65] F. Hong, X. Wang, Y. Yang, Y. Zong, Y. Zhang, and Z. Guo, "Wfid: Passive device-free human identification using wifi signal," in *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, 2016, pp. 47–56.
- [66] C. Shi, J. Liu, H. Liu, and Y. Chen, "Smart user authentication through actuation of daily activities leveraging wifi-enabled iot," in *Proceedings of the 18th ACM international symposium on mobile ad hoc networking and computing*, 2017, pp. 1–10.
- [67] Y. Zhang, Y. Zheng, G. Zhang, K. Qian, C. Qian, and Z. Yang, "Gaitsense: Towards ubiquitous gait-based human identification with wifi," *ACM Transactions on Sensor Networks (TOSN)*, vol. 18, no. 1, pp. 1–24, 2021.
- [68] X. Chen, H. Li, C. Zhou, X. Liu, D. Wu, and G. Dudek, "Fido: Ubiquitous fine-grained wifi-based localization for unlabelled users via domain adaptation," in *Proceedings of The Web Conference 2020*, 2020, pp. 23–33.
- [69] M. Ghifary, W. B. Kleijn, M. Zhang, D. Balduzzi, and W. Li, "Deep reconstruction-classification networks for unsupervised domain adaptation," in *Computer Vision–ECCV 2016: 14th European Conference*. Springer, 2016, pp. 597–613.
- [70] A. Zinys, B. van Berlo, and N. Meratnia, "A domain-independent generative adversarial network for activity recognition using wifi csi data," *Sensors*, vol. 21, no. 23, p. 7852, 2021.
- [71] Z. Shi, J. A. Zhang, R. Y. Xu, and Q. Cheng, "Environment-robust device-free human activity recognition with channel-state-information enhancement and one-shot learning," *IEEE Transactions on Mobile Computing*, vol. 21, no. 2, pp. 540–554, 2020.



Yubo Yan (Member, IEEE) received the BS, MS, and PhD degrees in communication and information system from the PLA University of Science and Technology, China, in 2006, 2011 and 2017 respectively. He is now an associate professor with the University of Science and Technology of China. His current research interests include Internet of Things, wireless networks, intelligent sensing, and mobile computing. He is a member of the IEEE Communications Society.



Dawei Yan received the BS degree in engineering from the School of Instrument Science and Electrical Engineering, Jilin University, China, in 2016. He is currently working toward the PhD degree with the University of Science and Technology of China. His current research interests include mobile computing and wireless sensing.



Panlong Yang (Senior Member, IEEE) received the BS, MS, and PhD degrees in communication and information system from the Nanjing Institute of Communication Engineering, China, in 1999, 2002, and 2005 respectively. He is now a professor with Nanjing University of Information Science & Technology. His research interests include wireless mesh networks, wireless sensor networks, and cognitive radio networks. He is a member of the IEEE Computer Society and ACM SIGMOBILE Society.



Xiang-Yang Li (Fellow, IEEE) received the bachelor's degree from the Department of Computer Science, Tsinghua University, China, in 1995, and the MS and PhD degrees from the Department of Computer Science, University of Illinois at Urbana-Champaign, in 2000 and 2001. He is a professor and executive dean with the School of Computer Science and Technology, USTC. He is an ACM fellow (2019) and ACM distinguished scientist (2014). He was a full professor with the Computer Science Department, IIT. His research interests include artificial intelligence of Things (AIOT), privacy and security of AIOT, and data sharing and trading. He and his students won several best paper awards, including MobiCom.



Fei Shang received the BS degree in electronic information engineering from the College of Information Science and Technology, Northwest University, China, in 2020. He is currently working toward the PhD degree with the University of Science and Technology of China. His current research interests include wireless sensing systems, wireless networks, and IoT Technology.



Feiyu Han received the BS degree from the School of Computer Science and Engineering, Nanjing University of Science and Technology, China, in 2019, and the MS and PhD degrees from the School of Computer Science and Technology, University of Science and Technology of China, in 2024. His current research interests include mobile/wearable computing, wireless sensing, and low-power communication.