MU-ID: Multi-user Identification Through Gaits Using Millimeter Wave Radios

Xin Yang*, Jian Liu†*, Yingying Chen*§, Xiaonan Guo‡, Yucheng Xie‡

*WINLAB, Rutgers University, New Brunswick, NJ, USA
†The University of Tennessee, Knoxville, Knoxville, TN, USA
‡Indiana University-Purdue University Indianapolis, Indianapolis, IN, USA
Email: xinyang@winlab.rutgers.edu, jliu@utk.edu, yingche@scarletmail.rutgers.edu, {xg6, yx11}@iupui.edu
§Corresponding Author

Abstract—Multi-user identification could facilitate various large-scale identity-based services such as access control, automatic surveillance system, and personalized services, etc. Although existing solutions can identify multiple users using cameras, such vision-based approaches usually raise serious privacy concerns and require the presence of line-of-sight. Differently, in this paper, we propose *MU-ID*, a gait-based multi-user identification system leveraging a single commercial off-the-shelf (COTS) millimeter-wave (mmWave) radar. Particularly, MU-ID takes as input frequency-modulated continuous-wave (FMCW) signals from the radar sensor. Through analyzing the mmWave signals in the range-Doppler domain, MU-ID examines the users' lower limb movements and captures their distinct gait patterns varying in terms of step length, duration, instantaneous lower limb velocity, and inter-lower limb distance, etc. Additionally, an effective spatial-temporal silhouette analysis is proposed to segment each user's walking steps. Then, the system identifies steps using a Convolutional Neural Network (CNN) classifier and further identifies the users in the area of interest. We implement MU-ID with the TI AWR1642BOOST mmWave sensor and conduct extensive experiments involving 10 people. The results show that MU-ID achieves up to 97% single-person identification accuracy, and over 92% identification accuracy for up to four people, while maintaining a low false positive rate.

I. INTRODUCTION

With the thriving growth of the smart Internet of Things (IoT), pervasive user identification brings huge convenience and changes to people's lifestyle. For instance, smart IoT systems could sense and recognize users to provide various personalized services. User identity-based authority management systems could keep the user's privacy and critical data protected. In addition, the prospering smart environment brings eager demands of simultaneous multiple people identification. The capability to recognize multiple users supports large-scale identity-based services, such as security surveillance and smart access control. However, traditional approaches relying on vision-based sensors (e.g., cameras) usually raise privacy fears and concerns. Thus, a solution that can achieve multiple people identification without compromising convenience and privacy is highly desirable.

Traditional user identification solutions are usually based on the knowledge of passwords (i.e., texts and graphical patterns) or physiological biometrics (e.g., voice, fingerprint, and iris) [4]. However, they all require users' intervention and are difficult to identify multiple users simultaneously. Moreover, camera-based approaches [12], [14], [19] based on the users' facial biometrics are proposed to identify single/multiple individuals without requiring extra effort. Nevertheless, these vision-based approaches record sensitive information such as user's appearance and face, which may violate user's privacy, especially when deployed in public areas. Recent studies rely on ambient WiFi signals, particularly channel state information (CSI), to identify individuals according to their distinct gait patterns [27], [29], [30] and the unique human physiological and behavioral characteristics inherited from their daily activities [22]. However, the aforementioned WiFi-based solutions all target on identifying a single user in a room as CSI is susceptible to the environmental changes, which restricts the ability to sense and identify multiple targets. In contrast, mmWave radar signals do not share these constraints due to shorter wavelength and stronger directivity, enabling superior multi-user sensing. For instance, Ahmad *et al.* [3] propose to monitor vital signs of multiple users simultaneously using an FMCW mmWave radar. Knudde *et al.* [11] use a mmWave radar to track multiple people's walking trajectory.

In this work, we take one step further and propose a passive multi-user identification system, called *MU-ID*, grounded on the unique human behavior characteristics (i.e., gait pattern) with a single COTS mmWave radar sensor. Comparing with other lower frequency RF signals (e.g., WiFi), the stronger directivity of mmWave radar makes it resistant to the environmental noises. Meanwhile, the shorter wavelength of mmWave signals could provide a much higher resolution of its sensing and imaging capability [28], making it an ideal candidate for capturing unique walking gait patterns of multiple users for identification. Notably, we derive the more stable lower-limb gait characteristics so that MU-ID could accommodate users' diverse upper-limb behaviors (e.g., carrying bags or holding smartphones) in their daily lives without compromising the identification accuracy. The proposed MU-ID has a broad range of application scenarios. For instance, the system could be deployed at the building entrance to identify staffs/residents for the purposes of access control and information statistics. Merchants such as hotels can deploy such a system at the front desk to establish gait profiles for guests and provide personalized services (e.g., turn on favorite TV channels, set preferred room brightness and air conditioner temperature)

before the guests arrive at their rooms.

In order to build such a robust and deployable multi-user identification system, we need to address many challenges in practice. Traditional WiFi-based gait recognition systems [27], [29], [30] identify the user based on the user's whole-body movements, which are also dependent on the user's irregular upper limbs behaviors (e.g., carrying a bag, and putting hands in the pockets). To mitigate the interferences associated with users' upper limbs, in this work, we use a compact low-power frequency-modulated continuous-wave (FMCW) radar with a narrow vertical field of view (FOV), operating in a high frequency band from $77GHz$ to $81GHz$, to only capture the users' fine-grained lower limb movements. Moreover, mmWave signals would be influenced by the movements of both the user's left leg and right leg, thus the lower limb of the user may be detected as two different moving targets. To solve this, we design a novel gait feature demonstrating the user's lower limb motion in the spatial-temporal domain leveraging the velocity and frequency response. The lower limb motion feature reflects multiple users' step lengths, step duration, instantaneous limb speeds, and inter-feet distances to perform multiple people identification. Additionally, our system also needs to get rid of the irrelevant signal reflected from static objects (e.g., wall, ceiling), which can be mistakenly treated as the user in the short time analysis. Hence, we denoise every frame of the radar signals in the range-Doppler domain to eliminate the interference from those static objects in the environment. The denoised user-specific gait profile is environment independent and can be applied to various environments without extra training. In the single-user profiling phase and multi-user identification phase, we first apply step segmentation approaches based on peak velocity and silhouette analysis to segment each walking step of users. For each step, the unique gait characteristics could be intuitively visualized through a heat map. Therefore, we use a trained deep Convolutional Neural Network (CNN) model, which has been broadly applied for image recognition, to identify step segments based on visual gait characteristics, and further identify the walking users in the area of interest.

The main contributions of our work are summarized as follows.

- To the best of our knowledge, MU-ID is the first work of performing multi-user identification based on the users' lower limb movements using only a single COTS mmWave radar sensor.
- We develop algorithms that have the capability to segment walking steps in both single-user and multi-user walking scenarios satisfying different application requirements.
- Our proposed system extracts unique environment independent user-specific features that integrate the walking step length, duration, instantaneous lower limb velocity, and inter-lower limb distance.
- Our proposed system is capable of mitigating the irrelevant interference from static objects in the environment and accurately identifying users in the area of interest grounded on a fine-tuned deep Convolutional Neural

Network (CNN).
- Extensive experiments involving 10 participants in two environments show that our system can achieve up to 92% accuracy only using 5 walking steps when identifying 4 users simultaneously while maintaining a low false positive rate.

II. Related Work

User identification systems relying on unique gait patterns can be broadly categorized as:

Sensor-based. Traditional methods mainly utilize various types of sensors, such as cameras, floor sensors (e.g., piezo-electric sensors), and wearable sensors to collect gait information for human identification. For instance, vision-based methods [8], [17], [26] use cameras to capture gait contour images and extract gait information to identify different people. However, these methods may raise privacy concerns and usually require the existence of a line-of-sight (LOS) condition. Ren *et al.* [21] design a user identification system leveraging gait patterns through run-time accelerometer measurements. Ruben *et al.* [25] propose a comparative assessment of footstep signals relying on floor sensors (i.e., piezoelectric sensors) and use the spatiotemporal information of the sensor signals as a new biometric for human identification. However, these identification schemes require users to either carry additional devices or install multiple sensors on the floor, which largely restricts their application scenarios.

WiFi-based. To overcome the above-mentioned weaknesses, some researchers seek WiFi-based solutions [22], [27], [29], [30] for gait-based user identification in a small group. For instance, WiWho [29] uses Channel State Information (CSI) of WiFi to extract distinguishing characteristics of walking gait for user identification. WifiU [27] identifies the user by extracting unique features from spectrograms of CSI and performing autocorrelation on the torso reflection to remove imperfection in spectrograms. However, these schemes require users to walk along a specific path, which is not very practical in many scenarios. Moreover, Cong *et al.* [22] utilize WiFi signals to capture unique human physiological and behavioral characteristics inherited from their daily activities, and further develop a deep-learning-based user authentication scheme to recognize each individual accurately. However, the aforementioned WiFI-based solutions can only identify a single user in a room as CSI is susceptible to the environmental perturbations collected from all directions.

Radar-based. Radar-based human sensing solutions attract considerable attention recently. Existing studies propose to use millimeter-wave radar sensors for gesture recognition [18], [31], indoor localization/tracking [2], [11] and vital signs monitoring [3], [13]. Additionally, some studies [1], [24] show that it is possible to leverage radio frequency signals to capture the human contour and perform human identification. Specifically, RF-Capture [1] reconstructs various human body parts and stitches them together through radio frequency signals. It shows the ability to capture a representative human figure and use it to distinguish between various users. Furthermore,
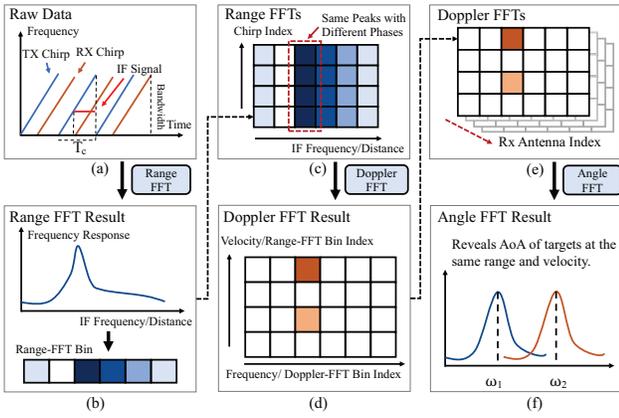
Fig. 1: Illustration of FMCW radar signal processing techniques.

Baptist *et al.* [24] propose to use the Micro-Doppler (MD) signatures retrieved from a low-power radar device to identify people based on their gait characteristics. However, there are limits in these schemas, because they either require the users to walk along a specific path or fail to recognize more than one person walking in the same room simultaneously, which is not very practical in real life.

Different from existing solutions, this paper proposes MU-ID, the first work that has the capability of recognizing multiple users simultaneously based on lower limb motion features via millimeter-wave radios. We exploit the unique gait characteristics (e.g., step length, the distance between two lower limbs, and instantaneous lower limb velocity) in the spatial-temporal domain through mmWave radio analysis. A CNN-based classifier is developed to extract users' gait characteristics and identify the users' walking steps, which is capable of facilitating many user identity-based applications and services.

III. Background of FMCW-based mmWave Radar

Typically, the walking process involves numerous body parts that create unique user-specific gait characteristics such as step distance, step duration and limb velocity [15], [16]. In this section, we introduce how the high resolution of a mmWave radar sensor could capture such unique gait characteristics by measuring the distance (e.g., step distance, limb position), velocity (e.g., the instantaneous velocity of limbs), and angle of arrival (AoA) necessary to distinguish multiple users.

A. FMCW-based Radar Technique

The radio signal travels at the speed of light, and thus, the time-of-flight (TOF), which is in a picosecond level, is extremely difficult to measure. In this work, we use FMCW-based radar technique that converts the challenge in directly measuring TOF to measuring frequency shift instead. In particular, the FMCW-based radar repeatedly transmits continuous chirp signals linearly sweeping through a frequency bandwidth B, and the duration of each chirp signal is T_c, as shown in Figure 1(a). Thus, the slope of the sweep is $S = \frac{B}{T_c}$. The received signal is the signal reflected from a target object, which is a delayed version of the transmitted signal. The radar

sensor then uses an integrated mixer generating the subtraction of the transmitted signal and the reflected signal, which is known as the intermediate frequency (IF) signal. The IF signal is separated by the unit of frame, which contains multiple chirps, to perform frequency domain analysis.

B. Range Measurement

The range information reflects the user's location, step distance, as well as the inter-lower limb distance. To measure the range of a target, we apply the FFT (i.e., range-FFT) on the time-domain IF signal. Particularly, when a user appears in the FOV, the strong frequency response of the human body produces a peak at a specific IF frequency representing the user's location. The distance between user and radar can be calculated as follows:

$$d = \frac{f_{IF} \cdot c \cdot T_c}{2 \cdot B} = \frac{f_{IF} \cdot c}{2 \cdot S}, \tag{1}$$

where the c is the speed of light and the frequency of the IF signal f_{IF} is derived by performing FFT (i.e., range-FFT) on a sliding window with a duration of T_c as shown in Figure 1(b). The result of range-FFT reveals the received frequency response at different ranges. Thanks to the centimeter-level range resolution, the range measurement precisely detects the position to the limb-level, which enables the measurement of the step length and the inter-lower limb distance.

C. Velocity Measurement

To further distinguish multiple users via the velocity of their limb movements, we exploit the phase of the range-FFT to derive the velocity in the Doppler domain. Particularly, the target with a moving speed v should have different phases in the two neighboring range-FFTs, which is associated with a movement distance vT_c during a chirp duration. Specifically, the relationship between the phase difference ω between neighboring range-FFTs and velocity v is shown as [23]:

$$v = \frac{\lambda \cdot \omega}{4\pi \cdot T_c}, \tag{2}$$

where λ is the wavelength. To distinguish multiple users using both range and velocity, another FFT (i.e., Doppler-FFT) is performed along the range-FFT columns as shown in Figure 1(c) to reveal the various speeds of multiple users at the same range. For instance, the two-colored bins in Figure 1(d) show two targets at the same range with various velocities.

D. Angle of Arrival Estimation

The range measurement only gives the distance of users in the LOS. Therefore, we use the AoA to depict the exact positions of users in a spatial Cartesian coordinate system. In particular, we exploit multiple receiving antennas of the FMCW radar to derive the AoA as follows [20]:

$$\theta = \sin^{-1}\left(\frac{\lambda \cdot \omega}{2\pi \cdot d_{inter-Rx}}\right), \tag{3}$$

where $d_{inter-Rx}$ is the distance between neighboring receiving antennas. To further distinguish between the overlap of multiple users in the range-Doppler domain, a third FFT (i.e., angle-FFT) is performed along all receiving channels. For instance, after applying angle-FFT on Figure 1(e), we can capture the
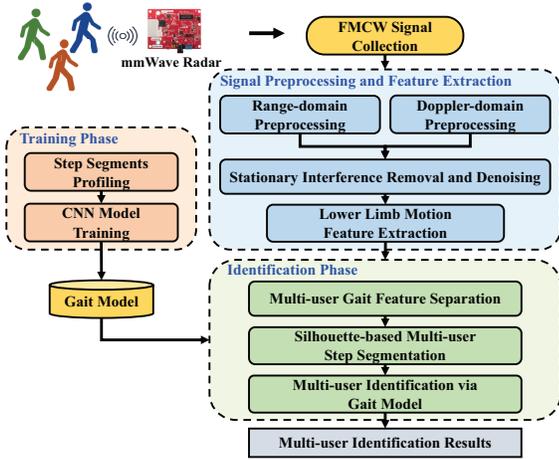
Fig. 2: Overview of MU-ID architecture.

(a) Range-Doppler domain signal.

(b) Stationary interference map.

(c) Interference removal result.

(d) Denoised range-Doppler signal.

Fig. 3: Illustration of interference removal and denoising.

AoA of two targets having the same range and velocity, as shown in Figure 1(f).

IV. CHALLENGES AND SYSTEM OVERVIEW

A. Challenges

To achieve multi-user identification using mmWave radios, a couple of challenges need to be addressed:

Capturing Gait Motions of Multiple Users. Traditional radar-based gait identification uses Micro-Doppler (MD) features [6] that can hardly reflect multi-users' gait patterns. Therefore, salient features should be designed and extracted to capture fine-grained lower limb motions for multi-user identification.

Separating Steps from Multiple Users. Commonly, multiple users may walk closely together (e.g., side by side, one after the other), merging the radar signal reflections of the individual lower limbs. Therefore, our system should be able to separate the walking gait of multiple users by step.

Eliminating the Stationary Interference. FMCW radar signals would be reflected by all the objects in the sensing field, which makes the signal analysis a challenging task due to the severe interferences caused by environmental noises and irrelevant reflections (e.g., reflections from wall and ceiling). Thus, it is essential to develop an interference removal algorithm that filters out the irrelevant signals so as to extract environment independent gait features for multi-user identification.

B. System Overview

The main architecture of MU-ID includes three phases - signal preprocessing and feature extraction phase, CNN-based gait model training phase, and multi-user identification phase, as shown in Figure 2. Particularly, the mmWave radar sensor takes the reflected mmWave signals as input and generates the IF signal from all receiving antennas (i.e., 4 receiving antennas). To derive the gait characteristics based on the range and velocity, we first perform range-FFT and Doppler-FFT on the received signals to reveal the distance and velocity of multiple users on a heat map. We then generate an interference
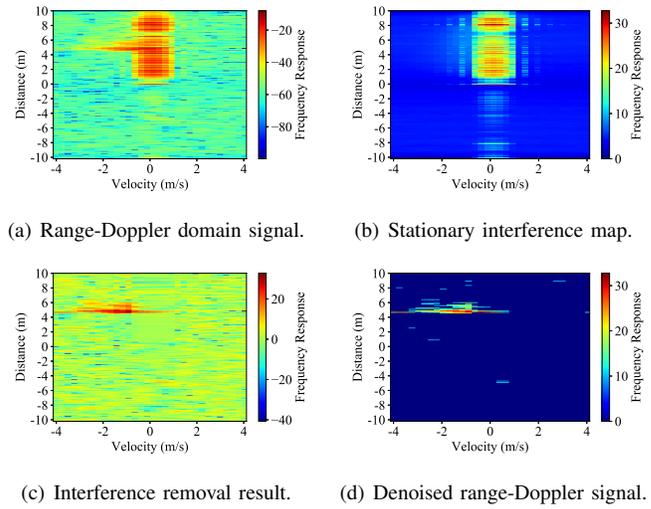
distribution map and use a high-pass filter to remove the interference caused by the stationary objects (e.g., wall and ceiling) and further calibrate the received signals. After this step, the range-Doppler map only retains the information associated with walking users. To demonstrate the instantaneous speed of users' lower limbs, we then derive the dominant velocity at every range leveraging the frequency response of the lower limb and the corresponding velocity. In the feature extraction step, we could obtain an intuitive heat map by integrating the velocity into the spatial-temporal dimension, which represents multiple users' distinct lower limb gait motions.

To construct the gait model for each user, we propose a single-user step segmentation method for the training stage profiling based on the dominant velocity peaks. The step segments reflect the unique gait pattern of each user. We then develop a CNN classifier to construct the user-specific environment independent gait model. In order to identify multiple users simultaneously through gaits, we derive a lower limb motion feature map that intuitively differentiates multiple users at distinct ranges. However, the lower limb motion map has limited abilities to describe multiple users at close proximity (e.g., walking abreast). To address this, a multi-user gait separation method is proposed to detach the gaits of multiple users in the range-Doppler domain by exploiting the AoA information to derive their spatial positions. To further segment individual steps, we use unsupervised learning techniques that automatically find the optimal strategy based on step silhouette. The steps are segmented by a fixed-size window and then identified through the gait model trained with a developed convolutional neural network.

V. SIGNAL PREPROCESSING AND FEATURE EXTRACTION

A. Range-Doppler-domain Preprocessing

To extract the user-specific gait pattern from the aspect of range and velocity, MU-ID converts the received raw FMCW signal into the range-Doppler domain. Particularly, MU-ID first performs chirp-wise range-FFT on the received signal to derive the range information of users. A strong frequency
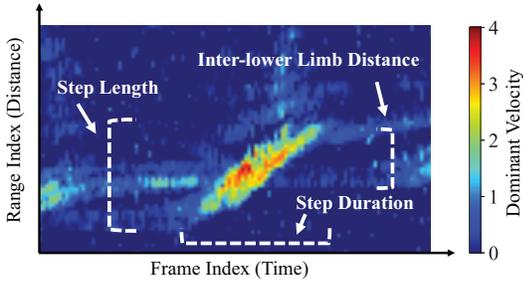
Fig. 4: Illustration of the lower limb gait feature.

Fig. 5: Correspondence between the actual step and lower limb feature map.

response F_p is caused by the reflected signal after bouncing off a reflector at a distance $\frac{F_p c}{2S}$ away from the radar, where S is the slope of the FMCW signal sweep. It is important to note that the signals would be reflected by many objects (e.g., walking users, wall, furniture, and ceiling) in the sensing field, each of which would result in strong frequency response. To further derive the various velocities of these reflectors, we then apply FFT on frequency-wise range-FFT signals in a frame, which is called Doppler-FFT. Figure 3(a) shows an example of the derived Doppler-FFT signal when a user is walking towards the radar sensor in a corridor. The x-axis (range-FFT bin) corresponds to the moving speed of the reflectors to the radar, in which zero indicates that the reflector is static. The y-axis (Doppler-FFT bin) corresponds to the distance of the reflectors to the radar.

B. Stationary Interference Removal and Denoising

As shown in Figure 3(a), the derived range-Doppler map contains the information of moving users, the stationary objects (e.g., wall, furniture, and ceiling), and environmental multi-path reflection effects [32]. In order to accurately capture the users' walking gait information, we need to eliminate the reflection effects from all of the stationary objects. Our radar sensor is configured to capture the fast-changing lower limb velocities at a high frame rate (i..e, $100 fps$). Therefore, for every frame, the frequency response of the user appears at a different range-Doppler position (i.e., distance in Figure 3(a)). In contrast, we find that the frequency response associated with the stationary objects in the range-Doppler map remains consistent over time. This motivates us to coarsely estimate the interference from the stationary objects by calculating the average frequency response in the range-Doppler domain, which can largely reduce the weights of the user-associated frequency response at each range-Doppler position.

Figure 3(b) shows the stationary interference distribution map derived from a three-second window, which contains 300 frames. To remove the interference, MU-ID subtracts the estimated stationary interference (e.g., Figure 3(b)) from each frame's range-Doppler domain frequency response (e.g., Figure 3(a)). The range-Doppler map after stationary interference removal is shown in Figure 3(c), which mainly contains the frequency response caused by walking users.

Additionally, we notice that the massive background noises accumulate severe deviations when processing the range-Doppler data. To eliminate the above impacts while preserving
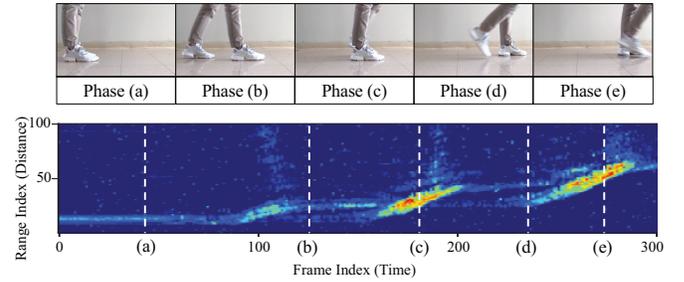

(a) Step of User A in Day 1.

(b) Step of User A in Day 2.

(c) Step of User B in Day 1.
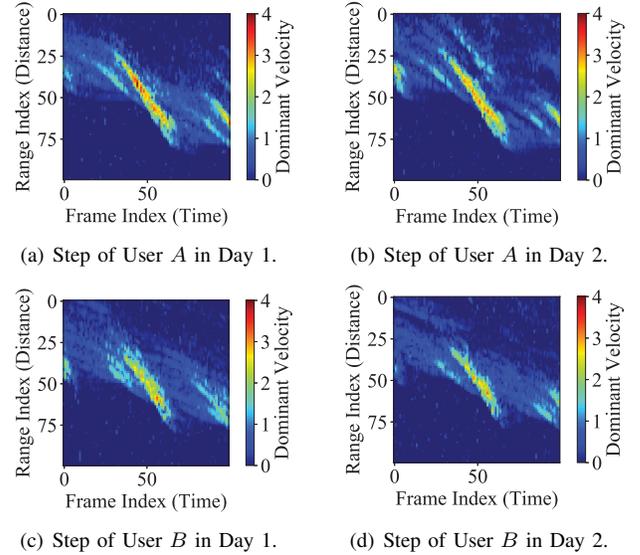

(d) Step of User B in Day 2.

Fig. 6: Comparison of the step segments from two users collected in different days.

the users' gait characteristics in the range-Doppler domain, we use a threshold-based high-pass filter described as below:

$$R_{(i,j,k)} = \begin{cases} R_{(i,j,k)}, & R_{(i,j,k)} \geq \tau, \\ 0, & R_{(i,j,k)} < \tau, \end{cases} \quad (4)$$

where $R_{(i,j,k)}$ is the range-Doppler domain frequency response at the k_{th} frame with range i and velocity j, and the threshold τ is empirically set to $10 dBFS$. The denoised range-Doppler map is shown as Figure 3(d), which has a much clearer user-related frequency response comparing to Figure 3(c).

C. Lower Limb Motion Feature Extraction

We observe that the limited velocity resolution resolves the velocity of one object at multiple adjacent bins as shown in Figure 3(d). Therefore, we use dominant velocity $\hat{V}_i$ to represent the user's lower limb velocity in every frame, which can be calculated as:

$$\hat{V}_i = \frac{\sum_{j=1}^{N_D}(\hat{R}_{(i,j,k)} V_j)}{N_D}, i \in [1, N_R], j \in [1, N_D]. \quad (5)$$

where $\hat{R}_{(i,j,k)}$ is the normalized frequency response; i, j, k stands for the range, velocity, and frame index, respectively; V_j is the velocity corresponding to the the frequency response $R_{(i,j,k)}$; N_R and N_D represent the number of range-FFT
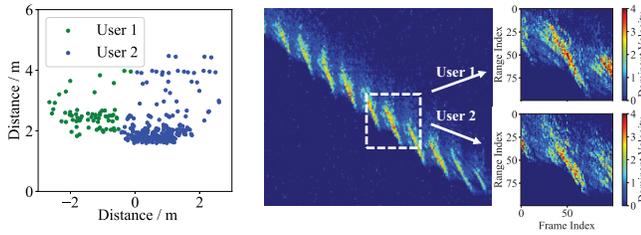
Fig. 7: Scatters of two abreast users in the spatial domain.

Fig. 8: Mixed steps of two abreast users (left), and separation results (right).

(a) Lower limb motion map of two users walking towards the radar.

(b) Step segmentation clustering result with step center marked as red.

Fig. 9: Illustration of multi-user step segmentation.

points and Doppler-FFT points, respectively. The dominant velocity integrates both the frequency response and velocity to compress a two-dimensional range-Doppler map into a one-dimensional array.

A sequence of the compressed range-velocity data is then arranged in the order of time, which is regarded as the lower limb motion feature map that shows user's gait characteristics (e.g., step length, duration, inter-lower limb distance and dominant lower limb velocity), as shown in Figure 4. To examine whether our proposed feature reflects the unique gait pattern of lower limb, we illustrate the correspondence between the actual gait process, which is captured by a camera, and the lower limb motion feature in Figure 5. The distance change over time reflects the step length and inter-lower limb distance; the step duration is indicated by the number of frames needed to finish the movement; the instantaneous velocity of the lower limb is represented by the dominant velocity and shown as the color map. Figure 6 shows the step segments from two users captured in two different days. The comparison shows that user A has a faster instantaneous lower limb velocity and a shorter step length comparing with user B; meanwhile the steps belonging to the same user shows high consistency. Thus, the spatial-temporal lower limb motion feature accurately reflects the user-specific lower limb gait pattern.

VI. MULTI-USER IDENTIFICATION

The crucial point in achieving multi-user identification is to detach the steps from simultaneously walking users. In this section, we introduce a lower limb motion separation technique and a silhouette-based multi-user step segmentation method leveraging unsupervised learning.

A. Lower Limb Motion Separation for Multiple Users

The spatial-temporal feature inherently differentiates multiple users at distinct ranges. However, such feature has an intrinsic limitation in distinguishing abreast walking users. Towards this end, we propose a lower limb gait feature separation approach exploiting the AoA. To detect the existence of abreast users, we empirically infer the range-Doppler domain frequency response summation of a single user as R_{single}. If the summation of frequency response at a specific range significantly exceeds the R_{single}, MU-ID confirms the existence of abreast users and perform the separation.

In order to separate the abreast users, we leverage users' distinct spatial positions. Specifically, we project the data
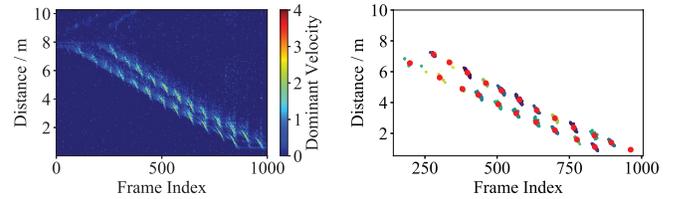
point $R_{(i,j,k)}$ in the k_{th} range-Doppler frame with range index i and velocity index j to a point $\hat{R}_{(i,j,k)}$ in the two-dimensional spatial Cartesian coordinate system, which takes the position of the radar as the origin. Leveraging $R_{(i,j,k)}$'s range D_i and angle of arrival $\omega_{(i,j,k)}$, the spatial coordinates $(X_{\hat{R}_{(i,j,k)}}, Y_{\hat{R}_{(i,j,k)}})$ of $\hat{R}_{(i,j,k)}$ is derived as below:

$$\begin{cases} X_{\hat{R}_{(i,j,k)}} = D_i \cdot \sin(\omega_{(i,j,k)}), \\ Y_{\hat{R}_{(i,j,k)}} = D_i \cdot \cos(\omega_{(i,j,k)}). \end{cases} \quad (6)$$

To differentiate the attribution of the scatters in the Cartesian coordinate system, we use the K-means clustering algorithm [7] to separate the scatters as shown in Figure 7. To determine the K (i.e., the number of users), MU-ID evaluates the value of K in $[1, K_{max}]$ using Calinski-Harabasz Index [5], where the K_{max} is estimated based on the average number of frequency response peaks. We then generate each user's individual gait feature in the range-Doppler domain by masking the frequency response from other users based on the spatial clustering results. Figure 8 shows the separation results of two abreast walking users that correctly identified by MU-ID.

B. Multi-user Step Segmentation based on Silhouette Analysis

In order to retain the user-specific step length and step duration features, we perform segmentation using a fixed-size window, which directly crops the step from the gait feature map. For most of the cases, the users' gait motions can be recognized from the spatial-temporal feature map, as shown in Figure 9(a). Therefore, the step segmentation needs to be capable of using for both single-user and multi-user scenarios. To achieve this, we use unsupervised learning technique to automatically detect the silhouette of steps and locate the step center for segmentation.

In order to make the silhouette of steps clearer, we first perform binarization to convert the lower limb motion map into a scatter plot. Particularly, the binarization uses a threshold η to set the data point $F_{(x,y)}$ in the feature map at frame x and distance y with a lower value as zero, and keep the one with a higher value as one. The threshold η can be selected through empirical studies. The binarized scatter plot is derived as:

$$\hat{F}_{(x,y)} = \begin{cases} 1, & F_{(x,y)} \geq \eta, \\ 0, & F_{(x,y)} < \eta. \end{cases} \quad (7)$$

Afterward, MU-ID categorizes the scatters using density-based spatial clustering of applications with noise (DBSCAN) algorithm. Based on the optimal clustering, we calculate the
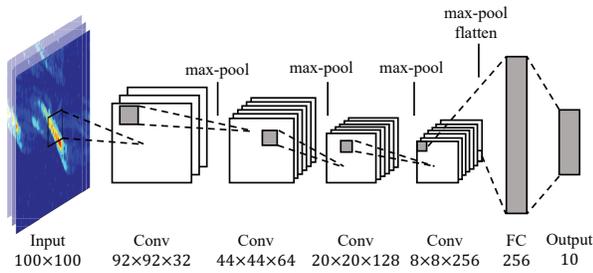
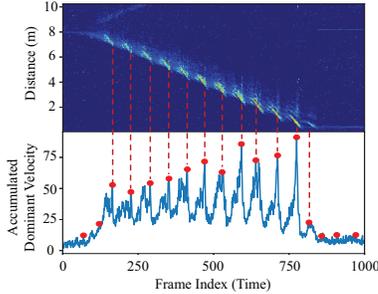Fig. 10: Illustration of the proposed CNN structure.

Fig. 11: Illustration of step segments profiling based on dominant velocity peaks.

coordinates of the cluster center to segment the steps using a fixed-size window. For a step cluster C contains N points $D_1, D_2, \cdots, D_N$, the coordinates of step center (X_{P_c}, Y_{P_c}) is derived as:

$$
\begin{cases}
X_{P_c} = \dfrac{\sum_{i=1}^{N} X_{D_i}}{N}, \\
Y_{P_c} = \dfrac{\sum_{i=1}^{N} Y_{D_i}}{N}.
\end{cases}
\tag{8}
$$

Figure 9(b) shows the step segmentation results of two users walking simultaneously with the step center marked as red.

VII. GAIT PROFILE CONSTRUCTION AND CNN-BASED MODEL TRAINING

A. Step Segments Profiling based on Dominant Velocity Peaks

To efficiently construct the user's gait profile at the training phase, we propose an intuitive and efficient single-user step segmentation method. Specifically, we locate each step from the lower limb gait map based on the observation that the area of the step center (i.e., the intersection where two lower limbs meet) possesses the highest dominant velocity. To derive the position of the step centers, MU-ID first calculates the accumulated dominant velocity as Figure 11 shows. The frame index of the step center corresponds to a peak in the accumulated dominant velocity. Through an efficient peak finding algorithm, MU-ID obtains the position of every step in the temporal axis first. Then a sliding window-based searching algorithm is used to locate the step in the spatial axis. To remove the fake peaks generated by the peak finding algorithm, MU-ID sets a minimum accumulated dominant velocity as a threshold to avoid invalid gait profiles.

B. CNN Model Training Through Gait Profiles

CNN is widely applied in the image recognition field for its ability in extracting fine-grained image patterns. To obtain the
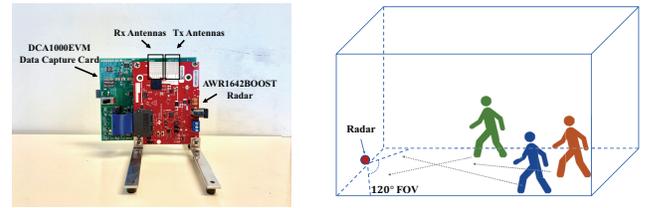
(a) AWR1642BOOST radar paired with DCA1000EVM data capture card.

(b) Illustration of the radar displacement and possible walk paths.

Fig. 12: Illustration of the experimental setup.

unique gait patterns from the two-dimensional step segments, we develop and fine-tune a deep CNN-based classifier. We illustrate the structure of MU-ID's CNN classifier in Figure 10. Particularly, the CNN takes as input the normalized $100 \times 100px$ step segments. Next, four convolutional layers varying in the number and size of convolution kernels extract the image-based gait patterns. Rectified Linear Unit (ReLU) is used as the activation function to reduce the vanishing gradient, followed by the max-pooling layer to downsize the intermediate output. The output of the last convolutional layer is flattened and processed by a fully connected layer with 256 neurons. Finally, a softmax layer outputs the probability distribution as the multi-user identification result.

VIII. PERFORMANCE EVALUATION

A. Experimental Methodology

1) Experimental Setup: We implement MU-ID on a COTS AWR1642BOOST [9] mmWave radar paired with a DCA1000EVM [10] data capture card as shown in Figure 12(a). The radar operates in a frequency band from $77GHz$ to $81GHz$ with a $4mm$ wavelength. The radar has two transmitting antennas and four receiving antennas lined up horizontally to give a sensing FOV of $120°$ in the E-plane and $30°$ in the H-plane, with $15°$ angle resolution. The radar is configured to sample 100 frames per second with 17 chirps per frame. The chirp slope is $18.83MHz/\mu s$, and each chirp contains 598 data samples. The ADC sampling rate is $3000k$ samples per second. The above configuration gives an $18m$ maximum detectable range with $4cm$ resolution and $8km/h$ maximum detectable velocity with $2km/h$ resolution. We displace the radar vertically on the floor pointing to the lower limbs. The evaluation is mainly performed in a $4m \times 8m$ lobby. We also test our system in another environment (i.e., a $1.6m \times 7m$ corridor) to evaluate the impact of environmental changes. The CNN-based classifier is based on TensorFlow framework, the model training and prediction is accomplished on a desktop server with GPU acceleration capability.

2) Data Collection: We recruit 10 participants aging from 21 to 34 with various heights from $164cm$ to $182cm$ and weights from $45kg$ to $74kg$. During the data collection process, we use a camera to record the ground truth and label the dataset. The radar is placed by the wall at the center of the lobby edge as demonstrated in Figure 12(b). Users walk from the opposite edge of the lobby and are allowed to walk from various starting points (within the FOV) towards the radar. The
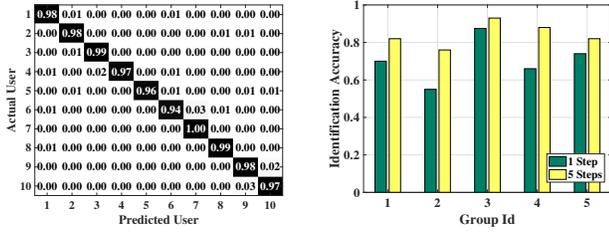
Fig. 13: Confusion matrix of single-user identification.

Fig. 14: Average accuracy of identifying two abreast users.

(a) Average accuracy of 1, 3, and 5-step identifiction.

(b) Average FP rate of 1, 3, and 5-step identifiction.

Fig. 15: Performance of multi-user identification.

radar records 10 seconds, therefore the path lengths fluctuate in the range from $5m$ to $7m$ based on the users' walk speeds.

Each participant is required to walk a total of eight times for the purposes of gait model construction and single-user identification. The data is collected under varying circumstances, including time of day, walk speed, and participant outfit. Unless mentioned otherwise, we split the dataset by a $6/4$ ratio, i.e., 60% of the data are used to construct gait models, and 40% are used to evaluate the performance.

For the multi-user scenario, ten participants are randomly split into groups with up to four members to walk simultaneously. Every group member is first required to keep a reasonable distance (e.g., $0.3m$) from other users to perform the experiments. Additionally, we evaluate the abreast walking scenario to demonstrate the performance under extreme cases. For every group size, we collect 20 sets of data. Apart from identifying multiple people based on a single step, we also evaluate the identification performance based on continuous three-step and five-step majority vote. Explicitly, when the majority vote ties, we specify the identification result to be the one with the highest average confidence. In total, we collected over 3,000 steps from the ten participants for evaluation.

3) Evaluation Metrics: To quantify the performance of our system, we define the following evaluation metrics:

Confusion Matrix. Each row in the confusion matrix is the user identity ground truth; each column represents the identity predicted by MU-ID. The entry of i_{th} row and j_{th} column shows the percentage of user i being identified as user j.

Identification Accuracy: The identification accuracy is defined as the probability of a user being correctly recognized by MU-ID.

False Positive Rate (FP rate): FP rate is defined as the ratio of other users (excluding user A) being mistakenly identified as user A.

B. Evaluation of Single-user Identification

We first examine the identification performance of MU-ID under the single-user scenario. For the single-user case, we identify the participants based on one step. Figure 13 shows the confusion matrix of single-user identification. Particularly, the average identification accuracy achieves 97.7% with a standard deviation of 1.6%. Among all the participants, the lowest identification accuracy achieves 94%. Additionally, the corresponding average FP rate is 0.2%, and the highest FP
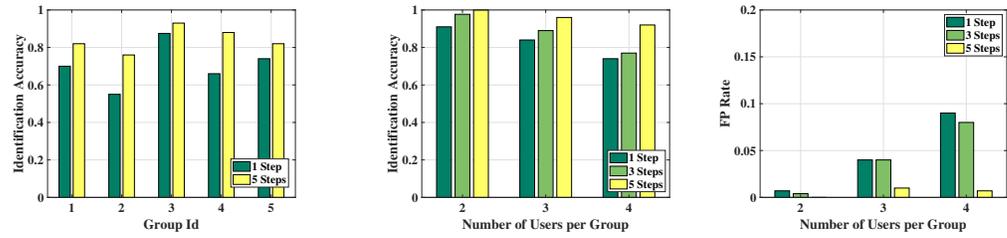
rate is 0.5%. This result validates that our system can identify individually walking user with high accuracy.

C. Evaluation of Multi-user Identification

We then evaluate the performance of multi-user identification. Figure 15(a) depicts the identification accuracy for up to four users walking simultaneously. Apart from the single-step performance, we demonstrate the performance when identifying three and five continuous steps.

For two-user identification, the average accuracy achieves 91% by identify a single step. When we apply majority vote on continuous three and five steps, the accuracy rises to 97.7% and 100% respectively. The FP rates for two-user identification are all below 1% as Figure 15(b) shows.

We then evaluate the three-user and four-user scenarios, where the average single-step accuracy descends to 84% and 74% respectively, and the average FP rates are around 4% and 9% respectively. The three-step majority vote slightly increases the accuracy to 89% and 77% for the three-user and four-user scenario respectively. When we identify continuous five steps, the accuracy dramatically rises to 96% for three-user identification and 92% for four-user identification. The above results indicate that MU-ID is capable of identifying multiple users simultaneously. Moreover, we find that the three-step majority vote is more sensitive to the misclassified steps, whereas the five-step majority vote accurately identifies multiple users.

We then evaluate the case when two users walk towards the radar side by side. We randomly combine two participants to walk abreast; a total of five combinations of participants are formed. Figure 14 shows the results of abreast user identification. When identifying a single step, MU-ID shows an average accuracy of 70%. With the help of the five-step majority vote, the accuracy increases to 84%. Considering that we only use a single mmWave radar and the angular resolution is limited $15°$, such accuracy of simultaneous multi-user identification is very encouraging.

D. Impact of Various Factors

1) Impact of Train/Test Split Ratio: We change the number of training data used for gait model construction to examine the performance of MU-ID under various train/test split ratios. To control other factors that could affect the identification accuracy, we evaluate the performance based on the single-user dataset. Each train/test split ratio is evaluated for five
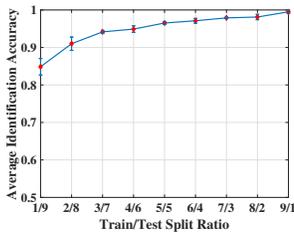
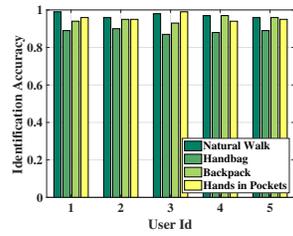Fig. 16: Error bar of various train/test split ratios.

Fig. 17: Identification accuracy with different upper limb behaviors.

(a) Identification accuracy under different environments.

(b) FP Rate under different environments.

Fig. 18: Impact of different environments.

times to calculate the average accuracy and deviation. We demonstrate the results in Figure 16. The results show that MU-ID has robust identification performance under various train/test split ratios. In particular, our model achieves over 96% average accuracy when the training set is no smaller than the test set. Even with the train/test ratio of 2/8, our system still retains an average accuracy above 90%. Meanwhile, the highest deviation is around 2.2% for the 1/9 train/test split ratio, which shows the performance of MU-ID is stable. Notice that even straight-line trajectories have variations across all rounds of tests, therefore our model is unlikely to be overfit. This result indicates that the features extracted by MU-ID show generality in represents the users' gait characteristics. Besides, MU-ID requires minimal training process to obtain the gait features.

2) Impact of Different Environments: We evaluate the identification performance under a non-training environment to validate that the gait feature is user-specific and insensitive to environmental changes. Specifically, we involve five participants to first train MU-ID with the dataset under the lobby environment. We then evaluate the identification performance under a new environment (i.e., a corridor having a size of $1.6m \times 7m$) based on the gait models derived under the lobby environment. Afterward, we exchange the training/testing environment. As Figure 18 shows, all of the involved participants can be recognized in the new environment. Specifically, when MU-ID identifies the participants under the corridor environment with gait models trained with the lobby dataset, the average accuracy and FP rate is 78% and 5% respectively. Exchanging the training/testing environment gives an average of 74% accuracy and 4% FP rate. Although User 2 and User 5 show below-average exception, most users achieve above 70% accuracy under a new environment. This result indicates that even without retraining the system, MU-ID still identifies most of the users' steps under a new environment. Notice that such results are based on identifying a single step, by applying the majority vote on continuous steps, higher identification accuracy is promising.

3) Impact of Upper Limb Behaviors: Finally, we study the identification performance when participants perform various upper limb behaviors to examine MU-ID's robustness under practical daily life scenarios. Five participants are required to perform three upper-limb behaviors while walking, i.e., 1) carrying a backpack, 2) grabbing a handbag with one hand,
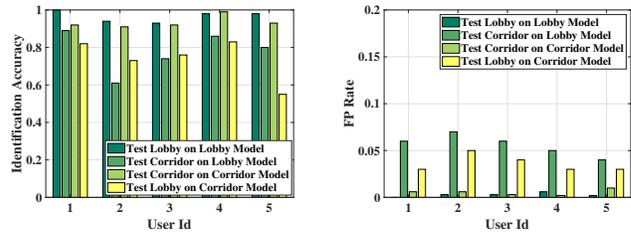
3) putting both hands in their trouser pockets. We present the identification results in Figure 17. The average accuracy for grabbing a handbag, carrying a backpack, and with hands in the pockets is 88%, 95%, and 96% respectively. Except for the handbag scenario, which shows a slight decrease in accuracy; the two other behaviors have consistent accuracy comparing with the 97% accuracy obtained when walking naturally. Specifically, when the participant grabs the bag with one hand, the handbag is close to the participant's lower limb. The signal reflections from the handbag can be captured within the radar's FOV. Therefore, the step segments contain both the handbag and user's lower limbs, which confuses the CNN classifier. We note that even with the interference from the handbag, the lowest identification accuracy still achieves 87%. The results reveal that MU-ID effectively verifies the lower limb movement and is robust to various upper limb behaviors. Such robustness is essential to the practical deployment in everyday life.

IX. CONCLUSION

In this paper, we study the feasibility of using mmWave radar to perform multi-user identification, which could facilitate various large-scale identity-based services. Towards this end, we design and implement MU-ID, a device-free gait-based multi-user identification system leveraging a single COTS mmWave radar. MU-ID exploits users' spatial-temporal lower limb motion features, including step length, duration, instantaneous lower limb velocity, and inter-lower limb distance to construct user-specific gait models. To simultaneously identify multiple users, we effectively separate the gait features from multiple users and perform silhouette-based analysis to segment individual steps. A CNN-based classification method is developed to extract gait features and perform multi-user gait identification. Extensive experiments involving ten participants show that MU-ID can achieve high identification accuracy and low false positive rate for both single and multiple people scenarios. In addition, a comprehensive study of different impacts (i.e., train/test split ratio, different environments, and upper limb behaviors) further confirms the robustness of our system.

X. ACKNOWLEDGMENT

REFERENCES

[1] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand. Capturing the human figure through a wall. ACM Transactions on Graphics (TOG), 34(6):219, 2015.

[2] F. Adib, Z. Kabelac, and D. Katabi. Multi-person localization via {RF} body reflections. In 12th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 15), pages 279–292, 2015.

[3] A. Ahmad, J. C. Roh, D. Wang, and A. Dubey. Vital signs monitoring of multiple people using a fmcw millimeter-wave sensor. In 2018 IEEE Radar Conference (RadarConf18), pages 1450–1455. IEEE, 2018.

[4] S. P. Banerjee and D. L. Woodard. Biometric authentication and identification using keystroke dynamics: A survey. Journal of Pattern Recognition Research, 7(1):116–139, 2012.

[5] T. Caliński and J. Harabasz. A dendrite method for cluster analysis. Communications in Statistics-theory and Methods, 3(1):1–27, 1974.

[6] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler. Micro-doppler effect in radar: phenomenon, model, and simulation study. IEEE Transactions on Aerospace and electronic systems, 42(1):2–21, 2006.

[7] J. A. Hartigan and M. A. Wong. Algorithm as 136: A k-means clustering algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(1):100–108, 1979.

[8] M. Hu, Y. Wang, Z. Zhang, D. Zhang, and J. J. Little. Incremental learning for video-based gait recognition with lbp flow. IEEE transactions on cybernetics, 43(1):77–89, 2012.

[9] T. I. Incorporated. Awr1642 single-chip 76-ghz to 81-ghz automotive radar sensor evaluation module. http://www.ti.com/tool/AWR1642BOOST, 2019 (accessed July 26, 2019).

[10] T. I. Incorporated. Real-time data-capture adapter for radar sensing evaluation module. http://www.ti.com/tool/DCA1000EVM, 2019 (accessed July 26, 2019).

[11] N. Knudde, B. Vandersmissen, K. Parashar, I. Couckuyt, A. Jalalvand, A. Bourdoux, W. De Neve, and T. Dhaene. Indoor tracking of multiple persons with a 77 ghz mimo fmcw radar. In 2017 European Radar Conference (EURAD), pages 61–64. IEEE, 2017.

[12] S. Z. Li, R. Chu, S. Liao, and L. Zhang. Illumination invariant face recognition using near-infrared images. IEEE Transactions on pattern analysis and machine intelligence, 29(4):627–639, 2007.

[13] M. Mercuri, I. R. Lorato, Y.-H. Liu, F. Wieringa, C. Van Hoof, and T. Torfs. Vital-sign monitoring and spatial tracking of multiple people using a contactless radar-based sensor. Nature Electronics, page 1, 2019.

[14] R. Min, J. Choi, G. Medioni, and J.-L. Dugelay. Real-time 3d face identification from a depth camera. In Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), pages 1739–1742. IEEE, 2012.

[15] M. P. Murray. Gait as a total pattern of movement: Including a bibliography on gait. American Journal of Physical Medicine & Rehabilitation, 46(1):290–333, 1967.

[16] M. P. Murray, A. B. Drought, and R. C. Kory. Walking patterns of normal men. JBJS, 46(2):335–360, 1964.

[17] M. Patel, P. J. Vandevord, H. Matthew, B. Wu, S. Desilva, and P. H. Wooley. Video-gait analysis of functional recovery of nerve repaired with chitosan nerve guides. Tissue engineering, 12(11):3189–3199, 2006.

[18] Z. Peng, C. Li, J.-M. Muñoz-Ferreras, and R. Gómez-García. An fmcw radar sensor for human gesture recognition in the presence of multiple targets. In 2017 First IEEE MTT-S International Microwave Bio Conference (IMBIOC), pages 1–3. IEEE, 2017.

[19] G. Porter and G. Doran. An anatomical and photographic technique for forensic facial identification. Forensic science international, 114(2):97–105, 2000.

[20] S. Rao. Introduction to mmwave sensing: Fmcw radars. Texas Instruments (TI) mmWave Training Series, 2017.

[21] Y. Ren, Y. Chen, M. C. Chuah, and J. Yang. User verification leveraging gait recognition for smartphone enabled mobile healthcare systems. IEEE Transactions on Mobile Computing, 14(9):1961–1974, 2014.

[22] 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, page 5. ACM, 2017.

[23] A. G. Stove. Linear fmcw radar techniques. In IEE Proceedings F (Radar and Signal Processing), volume 139, pages 343–350. IET, 1992.

[24] B. Vandersmissen, N. Knudde, A. Jalalvand, I. Couckuyt, A. Bourdoux, W. De Neve, and T. Dhaene. Indoor person identification using a low-power fmcw radar. IEEE Transactions on Geoscience and Remote Sensing, 56(7):3941–3952, 2018.

[25] R. Vera-Rodriguez, J. S. Mason, J. Fierrez, and J. Ortega-Garcia. Comparative analysis and fusion of spatiotemporal information for footstep recognition. IEEE transactions on pattern analysis and machine intelligence, 35(4):823–834, 2012.

[26] 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, 34(11):2164–2176, 2011.

[27] 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, pages 363–373. ACM, 2016.

[28] L. Yujiri, M. Shoucri, and P. Moffa. Passive millimeter wave imaging. IEEE microwave magazine, 4(3):39–50, 2003.

[29] Y. Zeng, P. H. Pathak, and P. Mohapatra. Wiwho: wifi-based person identification in smart spaces. In Proceedings of the 15th International Conference on Information Processing in Sensor Networks, page 4. IEEE Press, 2016.

[30] 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), pages 75–82. IEEE, 2016.

[31] Z. Zhang, Z. Tian, and M. Zhou. Latern: Dynamic continuous hand gesture recognition using fmcw radar sensor. IEEE Sensors Journal, 18(8):3278–3289, 2018.

[32] H. Zhao, R. Mayzus, S. Sun, M. Samimi, J. K. Schulz, Y. Azar, K. Wang, G. N. Wong, F. Gutierrez Jr, and T. S. Rappaport. 28 ghz millimeter wave cellular communication measurements for reflection and penetration loss in and around buildings in new york city. In ICC, pages 5163–5167, 2013.