IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 16 July 2024, accepted 9 August 2024, date of publication 13 August 2024, date of current version 26 August 2024.
*Digital Object Identifier 10.1109/ACCESS.2024.3443231*

RESEARCH ARTICLE

# Wi-Fi Signals for Passive Human Identification: A Study of Three Activities

**BAHA A. ALSAIFY[1], MAHMOUD M. ALMAZARI[1], RAMI ALAZRAI[2], (Member, IEEE), OMAR Y. AL-JARRAH[1], AND WALID AL ZIBIDEH[3], (Senior Member, IEEE)**
[1]Department of Network Engineering and Security, Jordan University of Science and Technology, Irbid 22110, Jordan
[2]Department of Computer Engineering, German Jordanian University, Amman 11180, Jordan
[3]Qualcomm Technologies Inc., San Diego, CA 92121, USA

Corresponding author: Baha A. Alsaify (baalsaify@just.edu.jo)

This work was supported by Jordan University of Science and Technology under Grant 20180032.
This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the IRB Office at King Abdullah University Hospital under Application No. 19/110/2017, and performed in line with the Declarations of Helsinki.

**ABSTRACT** This study proposes a passive human identification system based on wireless signals. The proposed system comprises three phases; preprocessing and standardization of recorded channel state information (CSI) signals and the extraction of relevant data using principal component analysis (PCA), transformation of the signals into feature vectors, and finally construction of classification models. To evaluate the proposed approach, we collected data from ten subjects in an indoor environment while performing a set of activities. The proposed approach was tested using three different activities: walking, sitting/standing, and picking up a pen, achieving subject identification accuracy of 97.3%, 99.75%, and 65.5%, respectively. The results suggest that activity-based identification systems can serve as an effective alternative to traditional methods such as passwords and smart cards.

**INDEX TERMS** Channel state information (CSI), gait recognition, human identification, Wi-Fi.

## I. INTRODUCTION

The recent advancements in communication technologies, particularly wireless communications, have facilitated the integration of wireless capabilities into various devices [1]. This presents researchers with the opportunity to utilize the abundance of wireless signals for various purposes, including the use of channel state information (CSI) to recognize and classify human activities [2], passive activity recognition [3], [4], [5], [6], subject identification [7], human-human interaction [8], [9], [10], and subject localization [11], enabled by advances in Machine Learning technologies. While CSI data has the potential to offer valuable insights into the movements and actions of individuals, it is sensitive to fluctuations in the environment.

Among the aforementioned applications, subject identification has been heavily researched with various techniques proposed in the literature. Historically, identification systems relied on secret information shared between the system and the subject participating in the identification process. However, such an approach has the risk of secret information being leaked or easily predicted if not chosen carefully. As a result, researchers have investigated alternative techniques such as fingerprint recognition [12], voiceprint analysis [13], and face recognition [14]. While these methods can be effective in determining a subject's identity, they do not utilize existing infrastructure in the same way that techniques utilizing overflowing wireless signals, such as CSI, do.

In this research, we present a passive human identification system that utilizes wireless communication between two devices to identify a human based on the activity being performed. The transmitter device sends Wi-Fi packets to the receiver, and any movement between the two devices will be reflected in the CSI values of the exchanged packets. By analyzing these CSI changes and extracting relevant features, we aim to determine the identity of the moving subject. To improve the accuracy of our system, we apply noise reduction and feature selection techniques to the

The associate editor coordinating the review of this manuscript and approving it for publication was Vicente Alarcon-Aquino.

recorded wireless signals. To validate the effectiveness of our approach, we conducted experiments using a dataset comprising ten subjects performing three different activities (walking, sitting down, and standing up from a chair, and picking up a pen from the ground). All experiments were conducted in the same environment but at different times to ensure the authenticity and reliability of the results. The proposed solution has several advantages compared to the traditional methods of subject identification. One of the main advantages is that it does not require any specific action from the subject, such as providing a password or scanning a fingerprint. Instead, it relies on the subject's natural movements, which are captured and analyzed using wireless communication technology. This makes the proposed solution a passive and non-intrusive method of subject identification. Another advantage is that it can be implemented using the existing infrastructure, namely, devices with wireless communication capabilities, without the need for any additional hardware or software. This makes the proposed solution a cost-effective and scalable solution. Finally, the proposed solution can be easily integrated into various applications, such as security systems, health monitoring systems, and smart homes.

The main contribution of our work can be summarized in the following:

1) All the experiments were performed using actual data acquired from real volunteering subjects. No simulations were performed;
2) A passive, non-intrusive method for subject identification is presented using CSI values, utilizing the existing infrastructure of wireless communication devices without the need for additional hardware or software, making it a cost-effective and scalable solution;
3) A methodology for noise reduction and feature selection to improve the accuracy of the proposed method is developed;
4) Different activities are tested for subject identification purposes; and
5) A new user identification approach based on new activities is presented.

The remainder of this paper is organized as follows; Section II, reviews the state-of-the-art methods while Section IV presents preliminaries Next, in Section IV, a description of the environment and the collected data is provided alongside the different data transformations that we apply to prepare the data for the identification procedure. The experiments performed and the obtained results are included in Section V while the final remarks are presented in Section VI titled as the conclusion.

## II. RELATED WORK

In this section, we review the existing approaches that focus on human identification using CSI signals. Most commonly used approaches in human identification systems [15] leverage the gait of subjects. Gait-based human identification systems can be categorized based on the utilized sensors into; camera-based, wearable-based, and Wi-Fi-based human identification systems.

In the first category, the gait of the subject is captured using cameras to identify subjects [16], [17], [18], [19], [20], [21]. Castro et al. [16] employed convolutional neural networks (CNNs) to learn the discriminative features, which are then fed into an SVM classifier to determine the identity of the subject, achieving comparable performance to the state-of-the-art. Li et al. [17] proposed an end-to-end gait recognition method that extracts the pose and shape features from an RGB gait sequence. Two gait datasets namely, OUMVLP [22] and CASIA-B [23], were used to evaluate the performance of the proposed system. Experiment results showed that the proposed system outperforms the existing approaches in terms of gait identification (assuming uncooperative subjects) and verification (assuming cooperative subjects).

In the second category, different wearable sensors, such as accelerometer, are used to capture the gait information of subject individuals [24], [25], [26], [27], [28]. The work of Sun et al. [24] uses a publicly available accelerometer dataset [29] to evaluate the performance of their proposed recognition system. The dataset was collected by attaching three accelerometer sensors to the body of the subject. Experiments' results suggest that to get the best results, the sensors should be placed in positions that are not affected by movements of body parts (e.g., not affected by hand movements). Although the researchers were able to achieve high accuracy, the proposed system restricts the subject's movements due to the attached sensors. Thus, they are not practical for everyday use. Zeng et al. [25] collected acceleration data [30] measured in the y-axis and z-axis using Wii Remotes attached to the subjects' bodies to train recognition models. The authors of [26] investigated the efficiency of using the built-in accelerometer in smartphones to record the gait data of subject individuals. However, the collected dataset was noisy as it was collected in a real-world uncontrolled environment. An equal error rate (EER) of 15.08% was achieved when trying to identify the different gait activities.

In the third category, Wi-Fi signals are used to capture human gait. Jakkala et al. [7] and Pokkunuru et al. [31] employed deep learning techniques to identify the subject individuals. In [31], the researchers used a deep CNN composed of 23 layers deep CNN to identify the subject individuals based on processed CSI measurements. The proposed approach achieved an average accuracy of 87.76%. Similarly, the work presented in [7] processes the collected CSI data by the use of a window-based denoising and normalization process. Deep CNN is then used to learn the discriminative features, achieving an average accuracy of 97.12%. Another work that focuses on utilizing the Wi-Fi CSI values is presented in [32]. In this work, the authors developed an approach for cross-state and cross-scene gait identification via unsupervised domain adaptive using Wi-Fi CSI values. To achieve the desired results, the authors implement a new data distribution metric, cross-attention distance, to achieve class-aware condition alignment at the class level. The experimental results demonstrate significant achievements, achieving a recognition accuracy of 94.61% in the cross-state gait recognition task.

## III. PRELIMINARY

Wi-Fi communication between two devices is based on Orthogonal Frequency Division Multiplexing (OFDM) technology, which is usually accompanied with Multi-input Multi-Output (MIMO) systems. The use of multiple transmitters and multiple receivers in any MIMO system establishes multiple streams between the sender and the receiver. The number of streams in any MIMO system is given by $N_{Tx} \times N_{Rx}$, where $N_{Tx}$ is the number of transmitting antennas and $N_{Rx}$ is the number of receiving antennas.

The received signal in each of the streams can be expressed as:

$$Y = H \times X + N, \tag{1}$$

where $X$ is the transmitted signal, $Y$ is the received signal, $H$ is the channel matrix, and $N$ is noise added to the signal while it propagates from the transmitter to the receiver.

The received OFDM signal is divided into different subcarriers, each of which can be quantified by the CSI values. Each CSI value is realized by two characteristics, namely the amplitude and phase, and can be expressed as follows:

$$CSI_i = |CSI_i|e^{j\angle CSI_i},\ i \in 1, \ldots, 30, \tag{2}$$

where $|CSI_i|$ and $\angle CSI_i$ represent the amplitude and phase of the CSI value associated with the $i^{th}$ subcarrier, respectively.

## IV. MATERIALS AND METHODS

In this section, we present the CSI dataset used to validate our proposed subject identification system. Moreover, we present the different stages comprised within our proposed system, including signal processing, feature extraction and selection, and building the classification models. Figure 1 provides a graphical illustration of our proposed system.

### A. DATA COLLECTION

To validate the performance of our proposed subject identification system, a set of experiments was designed to record the CSI values associated with the exchanged Wi-Fi signals. The experimental procedure was approved by the Institution Review Board (IRB) office at the Jordan University of Science and Technology under approval number (19/110/2017). Ten subjects participated in the data collection process. Before performing any of the experiments, each subject was asked to sign a consent form. The subjects were informed that their personal information would not be disclosed and that they had the right to stop participating in any of the experiments at any time if they chose to do so. Specifically, the subjects volunteered to perform a set of pre-explained activities between a pair of transmitting and receiving devices, each of them is equipped with "Intel Ultimate N Wi-Fi Link 5300" network card that was calibrated based on the settings provided in [33]. Using the tool described in [34], we were able to record the CSI values that describe the changes the subject movements exert on the exchanged Wi-Fi signals. It is worth mentioning that the CSI toolbox [34] uses the OFDM technique as a modulation scheme for wireless transmission channels. While collecting the data, we faced the issue of synchronizing the different subjects. All participating subjects must follow the same timing when performing the experiments to provide an unbiased dataset. To achieve this, we developed a sound-generating program to indicate to the participating subjects when to start the activity and when to finish it. Also, this program generates beeps to inform the subject of the start and finish of the sub-activities within each main activity. Another issue we faced was the existence of outliers in the collected dataset. These outliers have many sources. They might be introduced to the dataset following a hardware-related issue, or they can simply be introduced from the noise present in the environment since we are not performing the activities in a controlled environment. To combat these outliers, we are using the Hampel filter to remove them and smooth the data. More information of the Hampel filter can be found in Section IV-D1.

### B. ENVIRONMENT DESCRIPTION

All the experiments were conducted in an office room that has a length and width of 4.7 m and 4.7 m, respectively. Figure 1(A) shows a schematic diagram of the utilized office. The transmitter and receiver were set to be 3.7 m apart with no obstacles between them, thus all the experiments were performed in a Line-of-Sight (LoS) scenario. The subject performing the experiments was instructed to perform the activities in the area between the transmitter and the receiver.

### C. ACTIVITY DESCRIPTION

The volunteering subjects were asked to perform a predetermined set of activities. To synchronize the movements of the subjects, the recording device is programmed to emit a beeping sound to inform the subject of the end of the current sub-activity and the beginning of the next sub-activity. In this study, we investigated three different activities to determine the identity of a subject.

The first activity we are using to identify the subject is walking (i.e., Gait). A sample of a recorded walking activity signal is provided in Fig. 2, which shows that the walking activity comprises the following four activities: First, starting in front of the transmitter, the subject walks toward the receiver. The subject then turns when he/she reaches the receiver and walks back to the transmitter. Once the subject reaches the transmitter, he/she turns again, which concludes the current trial. By reviewing the literature, walking is the dominant activity used to determine the identity of the subject. In this work, in addition to walking, we are investigating two other activities. Specifically, we consider the activity of sitting on a chair, then standing up from

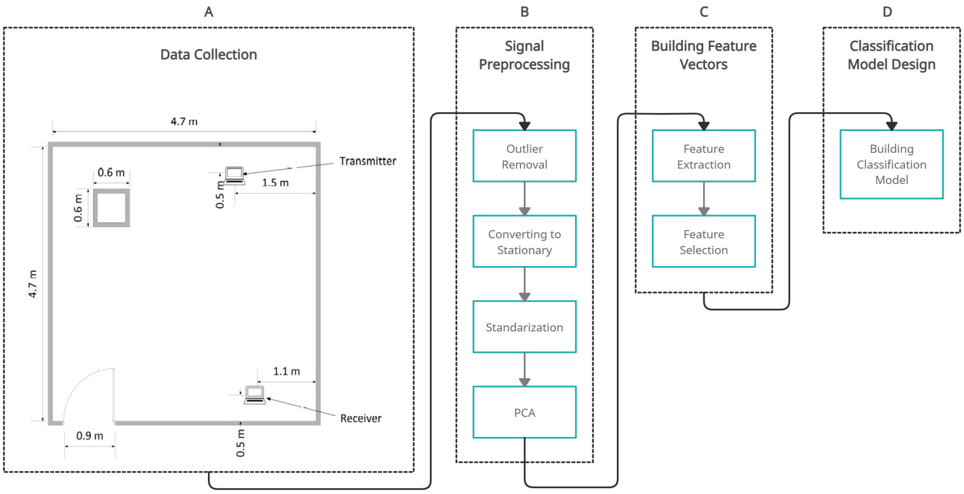



**FIGURE 1.** **The structure of our proposed subject identification system. (A) top view sketch of the data collection environment, (B) the main steps that comprise the signal processing stage, (C) the two main steps taken to transform the processed signal into a feature vector, and (D) the final stage of building the classification model.**

the said chair, and the activity of picking a pen up from the ground. A sample signal captured when performing the activity of sitting and standing up from the chair is provided in Fig. 2b. To perform the stand/sit activity, a chair was placed in the middle of the distance between the transmitter and the receiver. The subject was then instructed to start the activity sitting on that chair. When the subject hears a beep, they will stand up. Then, the second beep instructs the subject to stay still in their standing position. After that, the third beep informs the subject to sit on the chair and wait until the final beep is heard, which indicates the end of the activity. The last activity we are investigating is picking a pen from the ground. A sample signal recorded from picking a pen up from the ground is provided in Fig. 2c. The subjects were instructed to pick a pen from the ground within a specific time interval. The pen was placed in the middle of the distance between the transmitter and the receiver. The beginning and the end of the time interval are determined by hearing two beeps. Specifically, the first and second beeps indicate the beginning and end of the activity, respectively.

### D. PROPOSED SYSTEM

The purpose of this work is to provide a comparison between the effectiveness of different activities in determining the identity of the subject performing them. To that end, we propose an ML-based system capable of subject identity identification comprised of three phases, namely: signal processing, building representative feature vectors, and building the classification model.

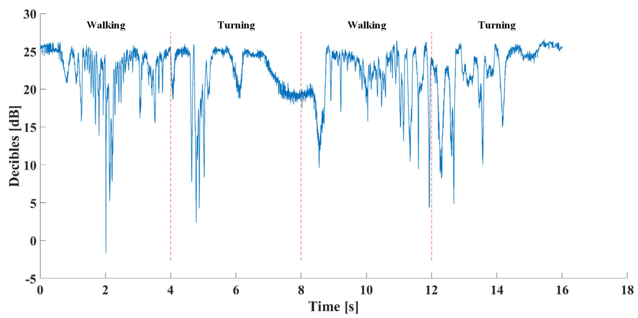
#### 1) PHASE1 (SIGNAL PROCESSING)

To prepare the data, we perform a multi-step signal transformation to remove the noise, converting the signal to a stationary signal, and standardizing it.

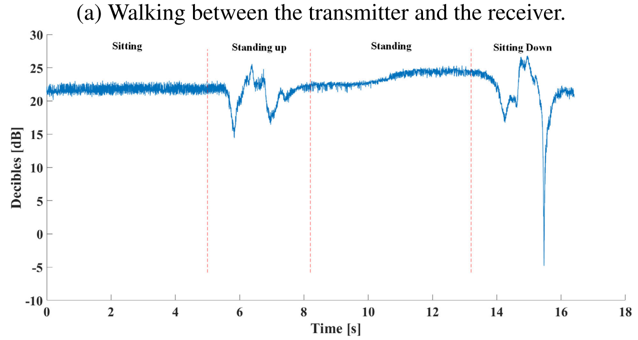
1) Removing outliers: outliers refer to sample points within a signal that are not following the natural progression of the signal. As suggested by other researchers [35], [36], [37], to remove these outliers, we applied the Hampel filter [38], which uses a sliding window with a fixed length. If the testing point deviates from the mean of the points in the window by more than three standard deviations, the point is removed and replaced by the mean of the points in the sliding window. The choice of using the Hampel filter can be attributed to many factors including: Using the Hampel filter does not distort the original data sequence since it replaces the detected outlier with the mean value of the defined window. Furthermore, the Hampel filter is adaptable to non-Gaussian noise thus in addition to Gaussian errors it can also detect non-Gaussian errors. Other justifications behind the use
2) Converting into a stationary signal: Stationary signals refer to a signal with constant or slowly changing statistical properties as described in [39]. To extract the statistical properties of a signal, it needs to be in stationary format. To determine whether a signal is stationary, a test called the Augmented Dickey-Fuller (ADF) test [40] is performed. If the hypothesis is that a unit root exists in the time-series signal, then the ADF test proves the hypothesis, and the signal is assumed to be non-stationary. On the other hand, if the test fails the hypothesis, then the time series signal is assumed to be stationary. If a signal is proven to be non-stationary, then we convert it to a stationary signal by using the differencing method. This method is applied as follows:

$$Y_{new} = Y_t - Y_{t-1}, \tag{3}$$

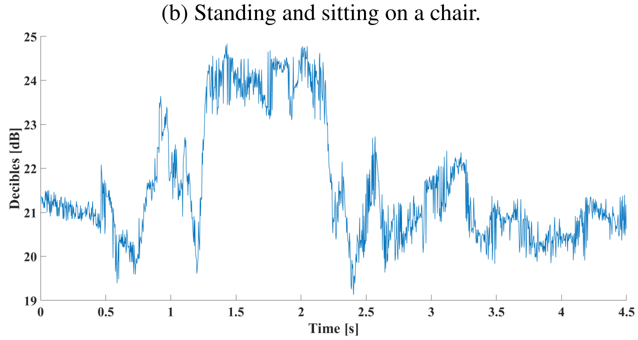
where $Y_t$ and $Y_{t-1}$ represent the value of the sample point within the signal at time indices $t$ and $t-1$,




(a) Walking between the transmitter and the receiver.




(b) Standing and sitting on a chair.




(c) Picking a pen from the ground.

**FIGURE 2. Samples of the signals for each of the three activities considered in this study. (a) illustrates a sample of the signal recorded for the walking activity, (b) illustrates a sample of the signal recorded for sitting down/standing up from a chair, and (c) illustrates the signal recorded when the subject picked up a pen from the ground.**

respectively, and $Y_{new}$ will replace the value of $Y$ at time $t$.

In addition to the ADF test we used to determine if the CSI time-series data is stationary or not, we reviewed the literature. We found several research articles such as [41] indicating that the CSI time-series data found in modern communication signals are non-stationary.

3) Standardization: The time-series signal is divided by the standard deviation of the signal to construct a signal with a standard deviation equal to 1, to ensure that the features have the same range, especially since all activities have the same time range. Thus, they are compared easily, and the performance of the classification model can be enhanced.
4) Dimensionality reduction: For each performed activity, we record data across 30 different subcarriers. However, building a classification model using data from all subcarriers is costly in terms of time, memory, and computation. Therefore, to reduce the amount of data while retaining data patterns, we use the Principal Component Analysis (PCA) technique [42].

   The different signal transformations are shown in Fig. 1(B). Moreover, Fig. 3 shows the effects of the different preprocessing transformations when applied to a sample signal.

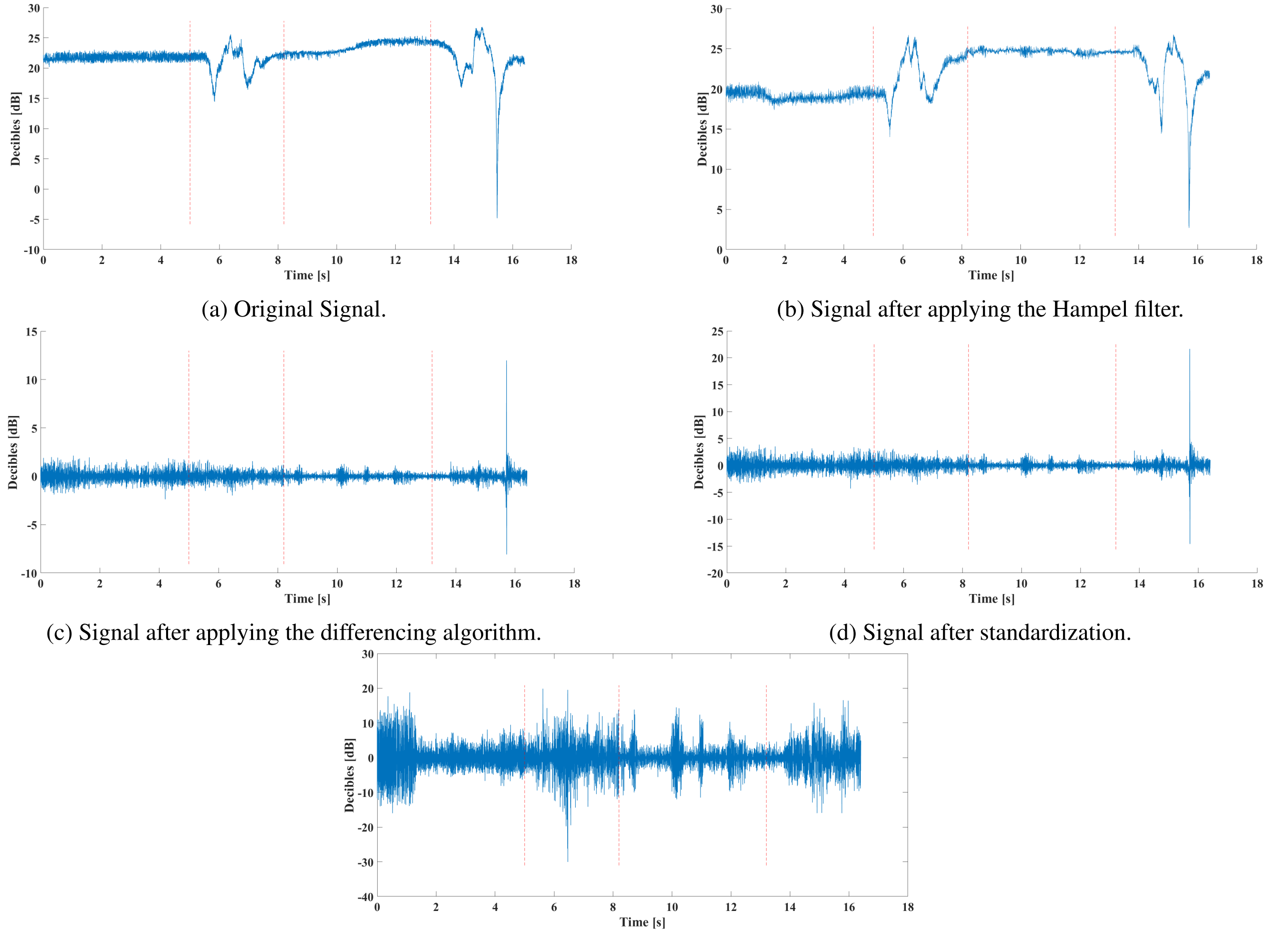
### 2) PHASE2 (BUILDING REPRESENTATIVE FEATURE VECTORS)

To build the classification models responsible for identifying the subject, several representative features must be extracted. These representative features, also known as feature vectors, are obtained by analyzing the already processed signals. A total of 565 features divided into 16 categories were extracted. Table 1 shows the features, their category, and a description of each presented feature.

Using all of these features to build the classification models, will have two primary effects; First, it will impose a burden on the processing using due to the large amount of data. Second, some features may collide with other features, degrading the performance of the classification model. To resolve these two problems, we select a subset of features using a feature selection algorithm as described in Algorithm 1. Specifically, we used the minimum redundancy maximum relevance (mRMR) method to reduce the number of features. Algorithm 1 shows the procedure by which we rank the extracted features. By examining Algorithm 1, only one pass over all the features is sufficient to determine the features that are higher than a certain threshold. Therefore, the complexity is of order $O(n)$ where $n$ is the number of available features. To determine the value of the threshold, we performed a series of preliminary experiments in which we investigated the effect of threshold value along with the number of used PCA components on the system's accuracy for each of the performed activities. From these experiments, the threshold value to distinguish between relevant and irrelevant features for each activity was determined.

### 3) PHASE3 (CLASSIFICATION MODEL)

To determine the identity of the subject, for each of the investigated activities a classification model was constructed. The first goal of the performed experiments is to determine the number of principal components and the number of features that result in the highest identification accuracy. For each component/features combination, a support vector machine (SVM) model was developed to identify the subject's identity performing the activities. Once the best combination is determined, we fine-tuned the SVM model by performing a grid search to determine the optimum SVM parameters, namely, the *gamma* parameter and the $C$ parameter.




(a) Original Signal.

(b) Signal after applying the Hampel filter.

(c) Signal after applying the differencing algorithm.

(d) Signal after standardization.

(e) Selected PCA component.

**FIGURE 3.** **The generated signals after applying each of the preprocessing transformations to a sample signal.**

---

**Algorithm 1** mRMR-Based Feature Selection

---

**input :** The principal components, each shown as a feature vector

**output:** The features sorted based on their importance

Normalize the features in the remaining feature vectors;
Apply the mRMR feature selection algorithm;
**for** *All features* **do**
  **if** *featureScore* > *Threshold* **then**
    FeatureToSelect = featureID (*mod featureCount*);
  **end**
**end**
Using the selected features, rebuild the feature matrix by excluding any unwanted features;

---

## V. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we evaluate the performance of our system and present a discussion of the results of our proposed subject identification system that are obtained for each of the three considered activities described in Section IV-C. We also compare the performance of our proposed system with other existing subject identification systems.

### A. EVALUATION PROCEDURES AND METRICS

To evaluate the performance of our proposed systems, we have developed four evaluation procedures, namely: Walking-based evaluation procedure, Sit/stand up-based evaluation procedure, Pick a pen from the ground-based evaluation procedure, and combined-based evaluation procedure.

In the first evaluation procedure, each subject is asked to walk between a signal transmitter and a signal receiver as explained in Section IV-C. To evaluate the performance of the proposed system, we arranged the walking data into four configurations. First, the whole 16-second data was used including the walking activity and the turning activities. Second, we removed the turning activity from the data and we used only the walking activity to train and test the classification models. For the third configuration, we only used the data gathered from when the subject walked from the transmitter to the receiver. Finally, for the

**TABLE 1.** Extracted features.

| Category | Feature | Description |
|---|---|---|
| Time series features | Mean | Several mean values were collected for: signal elements, linear weighted signal, power values in the signal, linear weighted power values, difference in adjacent elements, absolute difference in adjacent elements, power of difference in adjacent elements, and the absolute difference of differences. |
| | Skewness | Signal skewness |
| | Kurtosis | Signal kurtosis |
| | MAD | Signal mean absolute deviation |
| | Crossing rate | for 2, 1, 0, -1, -2 values |
| | RSSQ | Root sum of squares |
| | MAX | Maximum element in the signal |
| | MIN | Minimum element in the signal |
| | MAX -MIN | Difference between the max and min |
| | Quartile values | The first, second (also known as median), and the third quartile |
| | IQR | Inter quartile range (Q3-Q1) |
| | Normalized count of samples | For both the samples higher and lower than the mean of the signal. |
| | Element index | index of the first occurrence of the max and min signal elements |
| | cross-cumulants | second, third, and fourth order |
| | Number of peaks | based on a prominence value of 0.1 and 0.2 |
| | ARMA model parameters | Values for the AR, MA, and the model's constant |
| SNR | SNR | Signal to noise ratio |
| Frequency series features | Median frequency | The signal's median angular frequency |
| | Mean frequency | The mean angular frequency of; the entire signal, the first 10% of the signal, the next 10% of the signal, ..., the last 10% of the signal |
| | Bandwidth | The occupied bandwidth of the signal |
| | Power bandwidth | Compute the bandwidth of the part that is3-dB from the peak value |
| Linear predictor coefficients | LPC | The second, third, and fourth coefficients |
| Line spectral frequencies | LSF | Obtained from the prediction polynomial with coefficients obtained from the LPC |
| Distribution tests | Anderson-Darling test | The test result at a 0.05 level and the p-value |
| | chi-square test | The test result at a 0.05 level and the p-value |
| | Durbin-Watson test | The p-value |
| | Jarque-Bera test | The test result at a 0.05 level and the p-value |
| | Kolmogorov-Smirnov test | The test result at a 0.05 level and the p-value |
| | Lilliefors composite test | The test result at a 0.05 level and the p-value |
| | Runs test for randomness | The test result at a 0.05 level |
| Location tests | Wilcoxon rank sum test for equal medians | The test result at a 0.05 level and the p-value |
| | Wilcoxon rank sum test for zero medians | The test result at a 0.05 level and the p-value |
| | Sign test for zero medians | The test result at a 0.05 level and the p-value |
| | One-sample and paired-sample t-test | The test result at a 0.05 level and the p-value |
| | two-sample t-test | The test result at a 0.05 level and the p-value |
| Dispersion tests | Ansari-Bradley two-sample test | The test result at a 0.05 level and the p-value |
| | two-sample F test | The test result at a 0.05 level and the p-value |
| Stationary tests | Augmented Dickey-Fuller test | The test result at a 0.05 level and the p-value |
| | KPSS test | The test result at a 0.05 level and the p-value |
| | Leybourne-McCabe test | The test result at a 0.05 level and the p-value |
| | Phillips-Perron test | The test result at a 0.05 level and the p-value |
| | Variance ratio test | The test result at a 0.05 level and the p-value |
| | Paired integration/stationarity tests | The test result at a 0.05 level and the p-value for both the signal and the signal differences. |
| Correlation | Auto-correlation | Extract the time-series features from the auto-correlation sample signal |
| | Partial auto-correlation | Extract the time-series features from the partial auto-correlation sample signal |
| | Cross-correlation | Extract the time-series features from the cross-correlation between the first and second half of the signal |
| | Linear-correlation | Linear-correlation between the first and second halves of the signal |
| | Ljung-Box Q-test | The test result at a 0.05 level and the p-value |
| | Belsley collinearity diagnostics test | Strength of collinearity in the first and second halves of the signal |
| Causation test | Granger causality tests | The test result at a 0.05 level and the p-value |
| Heteroscedasticity | Engle test | The test result at a 0.05 level and the p-value |
| Cointegration | Engle-Granger cointegration test | The test result at a 0.05 level and the p-value |
| | Johansen cointegration test | The test result at a 0.05 level and the p-value for the first and second halves of the signal |
| Instantaneous frequency | Instantaneous frequency | Extract the time-series features from the temporal derivative of the oscillation phase divided by $2\pi$ |
| Spectral entropy | Spectral entropy | Extract the time-series features from the spectral entropy of the signal |
| Audio features | Spectral Centroid | Extract the time-series features from the spectral centroid of the signal |
| | Spectral Crest | Extract the time-series features from the spectral crest of the signal |
| | Spectral Decrease | Extract the time-series features from the spectral decrease of the signal |
| | Spectral Entropy | Extract the time-series features from the spectral entropy of the signal |
| | Spectral Flatness | Extract the time-series features from the spectral flatness of the signal |
| | Spectral Flux | Extract the time-series features from the spectral flux of the signal |
| | Spectral Kurtosis | Extract the time-series features from the spectral kurtosis of the signal |
| | Spectral Rolloff | Extract the time-series features from the spectral rolloff of the signal |
| | Spectral Skewness | Extract the time-series features from the spectral skewness of the signal |
| | Spectral Slop | Extract the time-series features from the spectral slop of the signal |
| | Spectral Spread | Extract the time-series features from the spectral spread of the signal |
| | MEL Spectrum | Extract the time-series features from the MEL spectrum of the signal |
| | Estimation of the fundamental frequency | Extract the time-series features from the fundamental frequency of the signal |
| | Signal loudness | Total loudness of the signal |

fourth configuration, only the data from when the subject moved from the receiver to the transmitter was used.

In the second evaluation procedure, data is organized into four configurations. First, the whole data capture was

**TABLE 2. The employed performance evaluation metrics. TP, TN, FP, and FN denote the true-positive, true-negative, false-positive, and false-negative values, respectively.**

| Metric | Equation | |
|---|---|---|
| Accuracy | $\frac{TP+TN}{TP+FN+TN+FP}$ | (4) |
| False-Positive Rate (FPR) | $\frac{FP}{TN+FP}$ | (5) |
| Precision | $\frac{TP}{TP+FP}$ | (6) |
| F1 Score | $(1+\beta^2)\times\frac{Precision\times TPR}{(\beta^2\times Precision)+TPR}$ | (7) |

used including the no-movement data portion, the sitting data portion, and the standing data portion. In the second configuration, we removed the no-movement data portion and used only data with sitting or standing activities. For the third configuration, we only used the sitting data portion, while the standing data portion was used for the fourth configuration.

In the third evaluation procedure, subjects are asked to pick a pen from the ground in any pose they are comfortable with.

The performance of our proposed system is evaluated by the following, shown in Table 2.

### *B. RESULTS OBTAINED FOR THE WALKING ACTIVITY*

To determine the effectiveness of the walking activity in determining the identity of the subject, we trained the classification model using four different configurations. First, we fed the whole 16-second recorded data into the SVM classifier and built the classification model based on this data. For the second configuration, we trained the classification model using the walking data alone without the turning activity that occurs at either the transmitter or the receiver. For the third and fourth configurations, we built the classification models based on the walking data from the transmitter to the receiver or by using the data from the subject walking back from the receiver to the transmitter.

For all the experiments, to determine the best number of principal components to use and the number of features (the threshold in Algorithm 1) to build the feature vector for classification, several components/features combinations were tested. All principal components between two and ten were tested. For each set of components, we varied the number of features. We started by selecting the top ten features and kept on increasing the number of features until we reached 500 features with increments of ten features. In other words, for each set of components, we tested the accuracy of the model with $\{10, 20, 30, \ldots, 500\}$ features. An illustration of the experiment where all walking data was taken into consideration is provided in Fig. 4.

It is worth mentioning that for each of the experiments, a 10-fold cross-validation technique was performed. In each of the folds, the model was trained on 90% of the data and tested with the remaining 10%. When we move to the next fold,

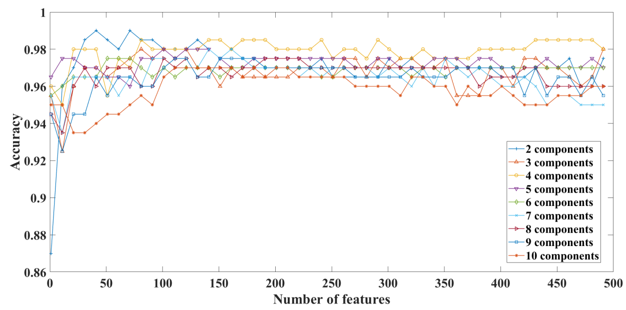



**FIGURE 4. PCA and feature selection for when the whole 16 seconds walking activity is used to train and test the model.**

another 10% of the data is selected for testing purposes and the remaining 90% is used for training the model.

#### 1) DATA FROM THE WHOLE EXPERIMENT

In this experiment, we fed the model with all the data acquired when the subject is walking between the transmitter and the receiver.

The results of the model's performance show that using only two components with only 50 features or 8.85% of the features results in an average accuracy of 99% across all folds.

#### 2) WALKING DATA WITHOUT TURNING

In this experiment, we removed the turning activity that occurs when changing the walking direction. In other words, we only used walking activity between the Wi-Fi transmitter and the Wi-Fi receiver that occurs within 0-4 seconds and 8-12 seconds of the experiment as shown in Fig. 2.

The results show that using only two components with 380 features or 67.25% of the features will return the best identification accuracy of 97.75%.

#### 3) WALKING FROM THE TRANSMITTER TO THE RECEIVER

In this experiment, we were only interested in the data capturing the subject walking from the Wi-Fi transmitter to the Wi-Fi receiver. Basically, from the timing diagram shown in Fig. 2, we extracted the data captured within the first four seconds of the experiment. As before, we tested multiple number of principal components, and for each number of components we tested a different number of features based on the features' score provided by the mRMR algorithm. The highest achieved accuracy was when 4 components were used with the top 280 features and it was equal to 98%.

#### 4) WALKING FROM THE RECEIVER TO THE TRANSMITTER

Similar to the previous experiment, only the data from when the subject was moving from the Wi-Fi receiver to the Wi-Fi transmitter was used to build the classification models. From the timing diagram shown in Fig. 2, we only used the data between the $8^{th}$ to the $12^{th}$ seconds of the experiment.

The results of the experiment show that an accuracy of 94.5% was achieved when 7 components were used with 180 features.

### C. RESULTS OBTAINED FOR THE SITTING DOWN AND STANDING UP FROM A CHAIR ACTIVITY

The second experiment we performed was to identify subjects based on the way they sat on a chair and stood from said chair. Similar to the walking activity, a timing sequence for this activity was developed and explained to the subject prior to the experiment. This timing diagram is provided in Fig. 2b. As illustrated in Fig. 2b, each experiment trial ran for 16-seconds. The subject started in a stationary position, then they stood from the chair, after that, they remained standing for 5 seconds and finally, they sat down on the chair, concluding the experiment trial. Each subject performed 20 trials of this activity.

To determine the best usage of the acquired data, we trained and tested the classification models using four different configurations. For the first configuration, we used the data collected for each trial without any alteration or reduction. For the second configuration, we removed the portion of the data in which the subject remained stationary. For the third configuration, only data associated with the sitting activity was used. As for the fourth configuration, only data associated with the standing activity was used. It should be mentioned that similarly to the walking experiment, a 10-fold cross-validation approach is used to determine the accuracy of the developed models. In each of the folds, we select 10% of the data to test the model and the remaining 90% to test the model. Across the folds, the selected data for testing the model will not be repeated. In other words, a data instance will be tested in only one fold. Additionally, in order to ascertain the optimal feature/PCA component pairing, akin to the approach taken for evaluating walking activity, we experimented with numerous combinations. For each combination, we recorded the identification accuracy. An illustration of this experiment is provided in Fig. 5

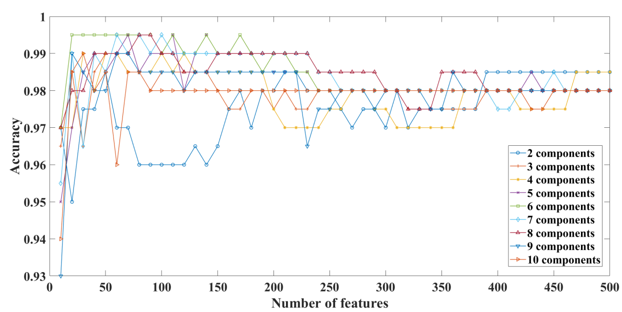



**FIGURE 5.** PCA and feature selection for when the entirety of the stand-sit experiment is included in the model for training and testing.

#### 1) DATA FROM THE WHOLE EXPERIMENT

In this test, we did not manipulate or divide the data in any way. The entirety of the data for each of the trials including the portion in which the subject was not moving and the portions where the subject was moving, were used to build and test the identification models. The length of each data instance used for model testing and training is equal to 16.4 seconds.

The results show that the highest achieved accuracy was equal to 99.5% when we are using the first 5 principal components with the 70 highest features according to the mRMR feature selection algorithm.

#### 2) STAND-SIT DATA WITHOUT THE NO MOVEMENT ACTIVITY

In this experiment, we removed the portion of the acquired data that represents a stationary subject. In other words, we are only considering the sections of the data in which the subject is sitting down on a chair or standing up from a chair.

The results show that an accuracy of 100% occurs when we are using 3 components with the highest 40 features according to the mRMR feature selection algorithm.

#### 3) DATA FROM ONLY THE SITTING ACTIVITY

In this experiment, only the sitting activity data was used to train the classification model. The results of this experiment show that using 8 principal components with the highest 100 features according to the mRMR algorithm will achieve a 100% identification accuracy.

#### 4) DATA FROM ONLY THE STANDING ACTIVITY

In this experiment, only the standing activity data was used to train the classification model. The results of this experiment show that using 8 principal components with the highest 60 features according to the mRMR algorithm will achieve a 99.5% identification accuracy.

### D. RESULTS OBTAINED FOR THE PICKING A PEN FROM THE GROUND ACTIVITY

In this experiment, we instructed the subject to pick a pen from the ground. The timing diagram for this experiment is provided in Fig. 2c which shows that each experiment trial lasted for 4.5 seconds in which the subjects started from a standing position, went down to pick a pen from the ground, and then they return to a standing position. Unlike the previous experiments where we divided the time frame into movement and no movement portions, in this experiment, we did not instruct the subject when to stop or when to move. The only limitation we imposed on the subjects was the time limitation to finish the experiment. Thus, only one configuration was used compared to the four configurations used for the previous two experiments. Similar to the other two activities, we tested many features/PCA component combinations. For each combination, the identification accuracy was logged. An illustration of this experiment is provided in Fig. 6

The results of this experiment showed that the highest accuracy of 64.5% was achieved when 8 principal components were used and each principal component was described using 100 features.

**TABLE 3.** Results summary.

| Activity | # Principal components | # Features | Accuracy | FPR | Precision | F1 score |
|---|---|---|---|---|---|---|
| Walking with turning activity | 2 | 50 | 99.00% | 0.112% | 99.04% | 99.00% |
| Walking without turning activity | 2 | 380 | 97.75% | 0.252% | 97.78% | 97.74% |
| Walking from Tx to Rx only | 4 | 280 | 98.00% | 0.224% | 98.05% | 97.98% |
| Walking from Rx to Tx only | 7 | 180 | 94.50% | 0.622% | 94.87% | 94.61% |
| Stand/Sit with stationary activity | 5 | 70 | 99.50% | 0.056% | 99.52% | 99.50% |
| Stand/Sit without stationary activity | 3 | 40 | 100.00% | 0.00% | 100.00% | 100.00% |
| Standing from a chair | 2 | 60 | 99.50% | 0.056% | 99.52% | 99.50% |
| Sitting on a chair | 8 | 100 | 100.00% | 0.00% | 100.00% | 100.00% |
| Picking a pen from the ground | 8 | 100 | 64.50% | 5.67% | 64.39% | 56.63% |

**TABLE 4.** Comparison with other existing approaches.

| Method | Number of Subjects | Number of Environments | Data Type | Activity | Classifier | Accuracy |
|---|---|---|---|---|---|---|
| Hong et al. [43], 2016 | 9 | 2 | CSI measurements | Gait | SVM | 92.8% |
| Xin et al. [44], 2016 | 9 | 1 | CSI measurements | Gait | KNN | 94.5% |
| Zhou et al. [45], 2019 | 15 | 2 | CSI measurements | Gait | softmax regression | 97.7% |
| Zhang et al. [46], 2020 | 20 | 1 | CSI measurements | Gait | DNN | 90.7% |
| Our proposed approach | 10 | 1 | CSI measurements | Walking | SVM | 97.31% |
| Our proposed approach | 10 | 1 | CSI measurements | Stand/Sit | SVM | 99.75% |
| Our proposed approach | 10 | 1 | CSI measurements | Picking a pen from the ground | SVM | 64.5% |

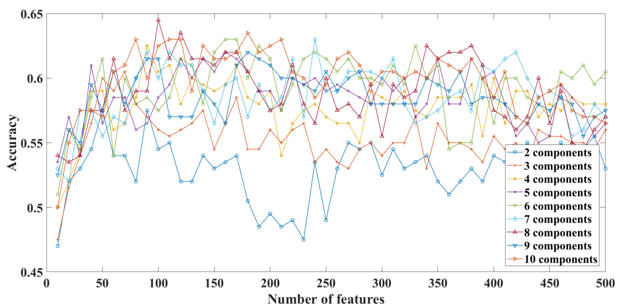



**FIGURE 6.** PCA and feature selection for when picking a pen from the ground activity is performed.

## E. DISCUSSION

Unlike other human identification systems that populate the literature, instead of relying on walking activity to determine the identity of a subject, in this work, we investigated other activities including standing from a chair, sitting on a chair, and picking a pen from the ground. From the evaluation of the performed experiments, several observations can be derived: First, using all available data does not necessarily result in the best performance. Using the developed system, we were able to achieve a 100% identification accuracy using only 3 principal components with only 40 features which is only 7% of the extracted features. Second, walking is not the only activity that can be used to determine the identity of the subject with high accuracy. In this work, we were able to achieve high identification accuracy by observing the way a subject stands and sits on a chair. In fact, the identification accuracy yielded from observing the way a subject walks was less than the accuracy yielded from observing the way a subject sits and stands from a chair. Third, the duration of actions directly influences subject identification accuracy, as demonstrated in our study. For instance, when subjects performed the task of picking up a pen from the ground, it took them approximately 4.5 seconds. In contrast, the action of sitting on a chair and then standing up took around 16 seconds. This short period affected the identification accuracy which was only 64.5%.

A summary of the results obtained by the performed experiments is provided in Table 3.

## F. PERFORMANCE COMPARISON

Our proposed system for subject identification outperforms other existing systems in terms of accuracy, as demonstrated in Table 4. This table compares our system with other systems based on the number of subjects, the number of environments, and the number of activities involved. It can be seen that our system achieved higher accuracy than the other systems.

# VI. CONCLUSION

Determining the identity of a person with high accuracy is the goal all authentication systems strive for. In this work, we propose a new methodology to determine a person's identity by extracting representative features from the recorded Wi-Fi signals influenced by the movements of said person. Several activities were investigated, including walking, standing and sitting on a chair, and picking a pen from the ground to determine which of them yields the highest identifying accuracy. The performed experiments reveal that walking, standing from a chair, and sitting on a chair activities result in high accuracy with slightly better performance for standing and sitting on a chair. A 100% accuracy was achieved when the standing and sitting on a chair activity was performed compared to 99% when the walking activity was performed, and 64.5% accuracy when picking a pen from the ground activity was performed.

In the future, we intend to extend the work presented in this manuscript by considering Non-Line-of-Sight (NLoS) scenarios for human identification, which can extend the range of the system and at the same time provide a more reliable identification system. Another aspect we plan on investigating is the effect the location of the subject has on the overall reliability of the system. Specifically, we intend to address the question: How does the proximity of the subject to the transmitter or the receiver affect the overall identification accuracy of the system? In addition to the previous question, we also plan to determine if it is possible to extract velocity and acceleration information from the collected dataset. If velocity and acceleration information were successfully collected, can they be used to build viable human identification models? Investigating the viability of other techniques is of importance to us. Thus, we plan on investigating several other techniques, specifically, deep learning techniques, such as DNN and LSTM on the available dataset(s) to determine if they perform better than classical classification techniques.

## DECLARATION OF INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

## REFERENCES

[1] J. Smith, "Wireless capabilities of modern devices," *J. Commun. Technol.*, vol. 25, no. 3, pp. 165–174, 2020.

[2] Y. Kim, J. Lee, and S. Park, "Classifying human activities using channel state information," *IEEE Access*, vol. 6, pp. 48671–48679, 2018.

[3] H. Yan, Y. Zhang, Y. Wang, and K. Xu, "WiAct: A passive WiFi-based human activity recognition system," *IEEE Sensors J.*, vol. 20, no. 1, pp. 296–305, Jan. 2020.

[4] B. A. Alsaify, M. M. Almazari, R. Alazrai, S. Alouneh, and M. I. Daoud, "A CSI-based multi-environment human activity recognition framework," *Appl. Sci.*, vol. 12, no. 2, p. 930, Jan. 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/2/930

[5] B. A. Alsaify, M. M. Almazari, R. Alazrai, and M. I. Daoud, "Exploiting Wi-Fi signals for human activity recognition," in *Proc. 12th Int. Conf. Inf. Commun. Syst. (ICICS)*, May 2021, pp. 245–250.

[6] C. Zhang and W. Jiao, "ImgFi: A high accuracy and lightweight human activity recognition framework using CSI image," *IEEE Sensors J.*, vol. 23, no. 18, pp. 21966–21977, Sep. 2023.

[7] K. Jakkala, A. Bhuya, Z. Sun, P. Wang, and Z. Cheng, "Deep CSI learning for gait biometric sensing and recognition," 2019, *arXiv:1902.02300*.

[8] R. Alazrai, M. Hababeh, B. A. Alsaify, M. Z. Ali, and M. I. Daoud, "An end-to-end deep learning framework for recognizing human-to-human interactions using Wi-Fi signals," *IEEE Access*, vol. 8, pp. 197695–197710, 2020.

[9] R. Alazrai, A. Awad, B. Alsaify, M. Hababeh, and M. I. Daoud, "A dataset for Wi-Fi-based human-to-human interaction recognition," *Data Brief*, vol. 31, Aug. 2020, Art. no. 105668.

[10] R. Alazrai, A. Awad, B. A. Alsaify, and M. I. Daoud, "A Wi-Fi-based approach for recognizing human-human interactions," in *Proc. 12th Int. Conf. Inf. Commun. Syst. (ICICS)*, May 2021, pp. 251–256.

[11] C.-H. Hsieh, J.-Y. Chen, and B.-H. Nien, "Deep learning-based indoor localization using received signal strength and channel state information," *IEEE Access*, vol. 7, pp. 33256–33267, 2019.

[12] K. Karu and A. K. Jain, "Fingerprint classification," *Pattern Recognit.*, vol. 29, no. 3, pp. 389–404, 1996.

[13] R. Brunelli and D. Falavigna, "Person identification using multiple cues," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 10, pp. 955–966, Oct. 1995.

[14] P. Tome, J. Fierrez, R. Vera-Rodriguez, and D. Ramos, "Identification using face regions: Application and assessment in forensic scenarios," *Forensic Sci. Int.*, vol. 233, nos. 1–3, pp. 75–83, Dec. 2013.

[15] C. Wan, L. Wang, and V. V. Phoha, "A survey on gait recognition," *ACM Comput. Surveys (CSUR)*, vol. 51, no. 5, pp. 1–35, 2018.

[16] F. M. Castro, M. J. Marín-Jiménez, N. Guil, and N. P. de la Blanca, "Automatic learning of gait signatures for people identification," in *Proc. Int. Work-Conf. Artif. Neural Netw.* Cham, Switzerland: Springer, 2017, pp. 257–270. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-59147-6_23#citeas

[17] X. Li, Y. Makihara, C. Xu, Y. Yagi, S. Yu, and M. Ren, "End-to-end model-based gait recognition," in *Proc. Asian Conf. Comput. Vis.*, 2020, pp. 1–17.

[18] Y. Wang, B. Du, Y. Shen, K. Wu, G. Zhao, J. Sun, and H. Wen, "EV-gait: Event-based robust gait recognition using dynamic vision sensors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 6351–6360.

[19] M. Deng and C. Wang, "Human gait recognition based on deterministic learning and data stream of Microsoft Kinect," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 12, pp. 3636–3645, Dec. 2019.

[20] N. Sadeghzadehyazdi, T. Batabyal, A. Glandon, N. Dhar, B. Familoni, K. Iftekharuddin, and S. T. Acton, "Using skeleton correction to improve flash LiDAR-based gait recognition," *Appl. Artif. Intell.*, vol. 36, no. 1, pp. 1–33, Dec. 2022.

[21] J. Chen, J. Fernandes, J. Ke, F. Lu, B. King, Y. H. Hu, and H. Jiang, "OptiGait: Gait monitoring using an ankle-worn stereo camera system," *IEEE Sensors J.*, vol. 24, no. 5, pp. 6888–6897, Mar. 2024.

[22] N. Takemura, Y. Makihara, D. Muramatsu, T. Echigo, and Y. Yagi, "On input/output architectures for convolutional neural network-based cross-view gait recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 9, pp. 2708–2719, Sep. 2019.

[23] S. Yu, D. Tan, and T. Tan, "A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, vol. 4, 2006, pp. 441–444.

[24] F. Sun, W. Zang, R. Gravina, G. Fortino, and Y. Li, "Gait-based identification for elderly users in wearable healthcare systems," *Inf. Fusion*, vol. 53, pp. 134–144, Jan. 2020.

[25] W. Zeng, J. Chen, C. Yuan, F. Liu, Q. Wang, and Y. Wang, "Accelerometer-based gait recognition via deterministic learning," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Jun. 2018, pp. 6280–6285.

[26] H. Alobaidi, N. Clarke, F. Li, and A. Alruban, "Real-world smartphone-based gait recognition," *Comput. Secur.*, vol. 113, Feb. 2022, Art. no. 102557.

[27] L. Qin, M. Guo, K. Zhou, J. Sun, X. Chen, and J. Qiu, "Gait recognition based on two-stream CNNs with multisensor progressive feature fusion," *IEEE Sensors J.*, vol. 24, no. 8, pp. 13676–13685, Apr. 2024.

[28] Q. Xiang, J. Wang, Y. Liu, S. Guo, and L. Liu, "Gait recognition and assistance parameter prediction determination based on kinematic information measured by inertial measurement units," *Bioengineering*, vol. 11, no. 3, p. 275, Mar. 2024.

[29] T. T. Ngo, Y. Makihara, H. Nagahara, Y. Mukaigawa, and Y. Yagi, "The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication," *Pattern Recognit.*, vol. 47, no. 1, pp. 228–237, Jan. 2014.

[30] Y. Zhang, G. Pan, K. Jia, M. Lu, Y. Wang, and Z. Wu, "Accelerometer-based gait recognition by sparse representation of signature points with clusters," *IEEE Trans. Cybern.*, vol. 45, no. 9, pp. 1864–1875, Sep. 2015.

[31] A. Pokkunuru, K. Jakkala, A. Bhuyan, P. Wang, and Z. Sun, "NeuralWave: Gait-based user identification through commodity WiFi and deep learning," in *Proc. 44th Annu. Conf. IEEE Ind. Electron. Soc.*, Oct. 2018, pp. 758–765.

[32] Y. Liang, W. Wu, H. Li, X. Chang, X. Chen, J. Peng, and P. Xu, "DCS-gait: A class-level domain adaptation approach for cross-scene and cross-state gait recognition using Wi-Fi CSI," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 2997–3007, 2024.

[33] B. A. Alsaify, M. M. Almazari, R. Alazrai, and M. I. Daoud, "A dataset for Wi-Fi-based human activity recognition in line-of-sight and non-line-of-sight indoor environments," *Data Brief*, vol. 33, Dec. 2020, Art. no. 106534.

[34] S. Radhakrishnan, S. Chinthamani, and K. Cheng, "The blackford Northbridge chipset for the Intel 5000," *IEEE Micro*, vol. 27, no. 2, pp. 22–33, Mar. 2007.

[35] I. O. Joudeh, A.-M. Cretu, R. B. Wallace, R. A. Goubran, A. Alkhalid, M. Allegue-Martinez, and F. Knoefel, ''WiFi channel state information-based recognition of sitting-down and standing-up activities,'' in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2019, pp. 1–6.
[36] B. Tan, Q. Chen, K. Chetty, K. Woodbridge, W. Li, and R. Piechocki, ''Exploiting WiFi channel state information for residential healthcare informatics,'' *IEEE Commun. Mag.*, vol. 56, no. 5, pp. 130–137, May 2018.
[37] H. A. H. Ali and S. Seytnazarov, ''Human walking direction detection using wireless signals, machine and deep learning algorithms,'' *Sensors*, vol. 23, no. 24, p. 9726, Dec. 2023.
[38] R. K. Pearson, Y. Neuvo, J. Astola, and M. Gabbouj, ''Generalized Hampel filters,'' *EURASIP J. Adv. Signal Process.*, vol. 2016, no. 1, pp. 1–18, Dec. 2016.
[39] G. P. Nason, ''Stationary and non-stationary time series,'' in *Statistics in Volcanology*, vol. 1. cLean, VA, USA: IAVCEI, 2006. [Online]. Available: https://pubs.geoscienceworld.org/gsl/books/edited-volume/1732/chapter-abstract/107597496/Stationary-and-non-stationary-time-series?redirectedFrom=fulltext
[40] G. Elliott, T. J. Rothenberg, and J. H. Stock, ''Efficient tests for an autoregressive unit root,'' *Econometrica*, vol. 64, no. 4, pp. 813–836, 1992.
[41] G. Liu, Z. Hu, L. Wang, H. Zhang, J. Xue, and M. Matthaiou, ''A hypernetwork based framework for non-stationary channel prediction,'' 2024, *arXiv:2401.08338*.
[42] S. Wold, K. Esbensen, and P. Geladi, ''Principal component analysis,'' *Chemometric Intell. Lab. Syst.*, vol. 2, nos. 1–3, pp. 37–52, 1987.
[43] F. Hong, X. Wang, Y. Yang, Y. Zong, Y. Zhang, and Z. Guo, ''WFID: Passive device-free human identification using WiFi signal,'' in *Proc. 13th Int. Conf. Mobile Ubiquitous Systems: Comput., Netw. Services*, Nov. 2016, pp. 47–56.
[44] T. Xin, B. Guo, Z. Wang, M. Li, Z. Yu, and X. Zhou, ''FreeSense: Indoor human identification with Wi-Fi signals,'' in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–7.
[45] Z. Zhou, C. Liu, X. Yu, C. Yang, P. Duan, and Y. Cao, ''Deep-WiID: WiFi-based contactless human identification via deep learning,'' in *Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Adv. Trusted Comput., Scalable Comput. Commun., Cloud Big Data Comput., Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Aug. 2019, pp. 877–884.
[46] 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 Things J.*, vol. 8, no. 9, pp. 7610–7624, May 2021.

**BAHA A. ALSAIFY** received the B.S. degree in computer engineering from Jordan University of Science and Technology, Irbid, Jordan, in 2007, and the M.S. and Ph.D. degrees in computer engineering from the University of Arkansas, Fayetteville, AR, USA, in 2009 and 2012, respectively. Then, he joined the Computer Engineering Department, Yarmouk University, Irbid. In 2015, he joined the Network Engineering and Security Department, Jordan University of Science and Technology, where he is currently an Associate Professor. His research interests include radio frequency identification (RFID), computer and network security, pattern recognition, human activity recognition, and human identification.

**MAHMOUD M. ALMAZARI** received the degree (Hons.) from Jordan University of Science and Technology (JUST). He is currently employed as a Software Programming Engineer with Atypon/Wiley, where he has excelled in Java programming and played a critical part in major projects including platform-wide improvements. He is a Committed Expert who specializes in network engineering, security, and software development. His academic career was highlighted by an outstanding graduation project on adaptive security operations. His scholastic and professional successes are complimented by a solid scientific background, which includes publications in areas, such as human activity identification utilizing Wi-Fi signals and RSA cryptanalysis. His dedication to perfection is further shown in his various honors, scholarships, and certificates.

**RAMI ALAZRAI** (Member, IEEE) received the Ph.D. degree in electrical and computer engineering from Purdue University, West Lafayette, IN, USA, in 2013. In June 2013, he became a Faculty Member with the School of Electrical Engineering and Information Technology, German Jordanian University (GJU), where he is holds the position of a Full Professor. In August 2022, he commenced a role as a Visiting Professor with the Department of Computer Science and Engineering, American University of Sharjah, United Arab Emirates. Concurrently, he serves as the Dean of the Scientific Research of GJU. His research interests include machine learning, biomedical image and signal processing, brain–computer interfaces, decoding mental imagery tasks through EEG signals, human activity recognition, and emotion recognition.

**OMAR Y. AL-JARRAH** received the B.Sc. degree in computer engineering from Yarmouk University, Irbid, Jordan, in 2005, the M.Sc. degree in computer engineering from The University of Sydney, Sydney, NSW, Australia, in 2008, and the Ph.D. degree in electrical and computer engineering from Khalifa University, Abu Dhabi, United Arab Emirates, in 2016. He is currently an Assistant Professor with Jordan University of Science and Technology, Irbid. He has more than 12 years of combined academic and industrial experience. His main research interests include machine learning, intrusion detection, big data analytics, autonomous and connected vehicles, unmanned aerial vehicles, and knowledge discovery in various applications.

**WALID AL ZIBIDEH** (Senior Member, IEEE) received the master's and Ph.D. degrees in electrical engineering from the University of Mississippi, Oxford, MS, USA, in 2010 and 2013, respectively. He is currently a Staff Engineer within Qualcomm's Operations Team. Since 2012, he has been with Qualcomm Technologies Inc., San Diego, CA, USA, where he held various technical positions in the field of power management IC (PMIC) test development.