IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 19 July 2024, accepted 31 July 2024, date of publication 5 August 2024, date of current version 22 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3438871

RESEARCH ARTICLE

Multi-User Human Activity Recognition Through Adaptive Location-Independent WiFi Signal Characteristics

FAHD ABUHOUREYAH[1], KOK SWEE SIM[2], (Senior Member, IEEE), AND YAN CHIEW WONG[1]

[1]Fakulti Teknologi dan Kejuruteraan Elektronik dan Komputer (FTKEK), Universiti Teknikal Malaysia Melaka (UTeM), Melaka 75450, Malaysia

[2]Faculty of Engineering and Technology (FET), Multimedia University, Melaka 75450, Malaysia

Corresponding author: Kok Swee Sim (sksbg2022@gmail.com)

This work was supported by Multimedia University Malaysia.

ABSTRACT In recent years, the remarkable advancement of WiFi sensing technologies has opened new frontiers in human activity recognition, enabling innovative solutions that transcend traditional methods and improve the capabilities of intelligent environments. Individual dynamic movements such as walking, sitting, standing, and running, as well as more complex interactions such as sports activities, are all examples of human activity. WiFi sensing has emerged as a powerful tool for human activity recognition; however, certain restrictions persist, especially when sensing activities involving multiple users across different locations. These limitations highlight the need for innovative techniques to address the intricacies of multi-user scenarios and environmental effects, ensuring the robustness and accuracy of WiFi-based sensing systems. To address multi-user effects in WiFi signals, we propose a few layering LSTM deep learning models with Raspberry Pi for edge computing solutions. The method leverages the decomposition of Channel State Information (CSI) signals through Independent Component Analysis (ICA) and Continuous Wavelet Transform (CWT). The integration of signal decomposition and deep learning holds promise for advancing WiFi sensing systems' accuracy, reliability, and real-time capabilities in complex environments and multi-user scenarios. Experimental findings prove the system's ability to handle complex activities with high classification accuracy. Furthermore, the system displays a remarkable ability to classify complex activities. By leveraging the power of deep learning, the model learns intricate patterns and relationships within the decomposed CSI signals, enabling it to distinguish between diverse activities with high accuracy.

INDEX TERMS WiFi sensing, multi-user, human activity recognition, channel state information, independent component analysis.

I. INTRODUCTION

The widespread deployment of WiFi infrastructure has resulted in the pervasive availability of WiFi signals, offering extensive convenience on a large scale while reducing costs. Simultaneously, the inherent non-intrusiveness of WiFi devices enhances the user experience by eliminating the need for wearable devices [1]. Currently, there is a growing interest in monitoring human motion perception, which holds significant potential for identifying human movement activities in various domains, including motion analysis, auxiliary medicine, virtual reality (VR) [2], and human-computer interaction [3]. Recent methods have facilitated innovative solutions that surpass traditional limitations and enhance the capabilities of intelligent environments. However, despite the demonstrated proficiency in recognising and detecting various movements within the realm of action recognition and counting, as evidenced in the existing literature [4], [5], [6], these advancements have been limited to single-person scenarios.

CSI monitoring using WiFi sensing offers an advantage in addressing privacy concerns compared to traditional vision-based monitoring systems [7]. CSI-based systems measure the changes in WiFi signal patterns caused by movements and interactions within an environment, enabling monitoring while maintaining a higher degree of anonymity.

The associate editor coordinating the review of this manuscript and approving it for publication was Ronald Chang.

This approach minimizes the risk of privacy invasion as it does not produce or rely on images or video data, thereby offering a less intrusive and more secure alternative for applications [8]. Moreover, multi-user sensing facilitates a deeper understanding of human behaviour and interactions, leading to insights that can inform decision-making processes and develop better intelligent and responsive systems. In addition, there are persistent challenges in recognising and comprehending human activities, especially in scenarios that involve multiple users across different locations [9]. Fig. 1 illustrates how line-of-sight (LOS) and non-line-of-sight (NLOS) conditions influence Channel State Information (CSI) signals, emphasising their location-dependent and environmentally influenced nature. The arrows and labels in the figure illustrate the variations, reflections, and distortions in the signals, thus highlighting the dynamic characteristics of wireless communication in complex and variable environments [10].

FIGURE 1. **Multi-users effects on CSI signals in wireless sensing.**

Existing methods face challenges in addressing the complexities of human activities, such as subtle gestures and dynamic movements, particularly in scenarios involving multiple individuals and using WiFi sensing technology. Additionally, environmental factors like interference and dynamic changes distort signals, such as frequency, amplitude, phase, and the precise interpretation of desired signals. The available techniques necessitate training to recognise constant data for activities that can be performed in multiple locations [6]. Additionally, these techniques are limited to individual person sensing capabilities due to the complexity of multi-person sensing [11], [13], [14]. These challenges highlight WiFi sensing systems' deep-rooted difficulties in sensing and interpreting activities involving multiple individuals.

These highlighted challenges require new and innovative approaches to achieve the stability and precision of WiFi-based sensing in multi-user scenarios, such as [7] and [15]. This work provides an innovative approach integrating activity recognition to address this intricate challenge. The proposed method identifies and quantifies multiple users within a given scene, independent of the specific location. The system monitoring CSI employs advanced signal processing techniques, including Independent Component Analysis (ICA) and Continuous Wavelet Transform (CWT). The method utilises path and propagation delay information to establish a threshold to retain only relevant path data that traverses the target body. Integrating a deep learning paradigm augments the system's capacity to discern intricate patterns within the decomposed CSI signals. The present research aims to develop a precise system capable of recognising and classifying human activities solely through ubiquitous WiFi infrastructure, eschewing the need for specialised sensing hardware. The key contributions of this work are are listed as follows:

1) Implementation of a CSI-based sensing to enable continuous human action recognition for multiple individuals utilising a WiFi system. This approach enables to reduce the influence of multi-user effects while acquiring path and amplitude information to improve activity classification accuracy.
2) The proposed model uses Independent Component Analysis ICA and Continuous Wavelet Transform CWT along with few layers Long-Short-Term-Memory LSTM algorithm to achieve location-independent multi-user sensing using a minimal dataset and less training effort, and achieves higher accuracy classification.
3) The experimental analysis explores the retention of action characteristics after decomposition and demonstrates the proposed method's effectiveness in detecting and counting activities with multiple users in different scenarios to evaluate the model across real-world scenarios and environments.

The proposed method implements the intricate challenges of diverse activities in a multi-user WiFi-based sensing scenario. It enables us to distinguish between different sources of signals by utilising the ICA method to analyse the motion of multiple individuals. This approach allows for precise identification and tracking of individual actions. Furthermore, the tensor decomposition and implementation methods leverage the CWT technique to capture the amplitude of information action. Required experiments were conducted to analyse the capability of the proposed method and focus on preserving action features for recognition after decomposition. The results of these experiments show the proposed approach's efficacy and potential.

The following section investigates the existing landscape of multi-person action recognition literature. Subsequently, section three introduces the proposed methodology, employing decomposition and pioneering algorithms. Section four validates the multi-user Human Activity Recognition (HAR) method through various experiments. The analysis engage in a detailed discussion of results and their implications, drawing comparisons with established methodologies and suggesting future directions for scholarly inquiry. The conclusion offers a conclusive summary, culminating the sequential discourse and emphasising the distinctive contributions of the proposed method.

II. RELATED WORKS

Research on HAR in multi-user environments has gained attention due to its growing applications in diverse domains such as healthcare [16], senior citizen monitoring [17], and smart environments [18]. The growing ubiquity of sensor-rich environments, pervasive computing technologies, and the increasing integration of intelligent systemscan be found into our daily lives. The imperative to discern and interpret human activities in settings with multiple users stems from its critical role in augmenting the capabilities of these intelligent systems. In healthcare, precise monitoring and recognition of activities offer unprecedented patient care and rehabilitation opportunities [16]. The complexity of multi-user environments introduces several challenges that necessitate advanced techniques [19], [20]. Occlusion, multi-user dynamic effects, and interruption of signal interactions pose problems for traditional wireless sensing systems.

The challenge of HAR for multi-users resides in the complexities arising from multiple individuals' simultaneous presence and interactions within a shared space [21]. The available HAR methods are primarily designed for single-user scenarios and confront a substantial hurdle when faced with the layered complexities inherent in multi-user environments [22], [23]. The core of this challenge lies in the need for more conventional HAR systems to discriminate between the activities of different users, ultimately resulting in a pronounced diminishment of accuracy and reliability [24]. The intricate interplay of human movements and the potential for occlusion and overlapping actions pose a formidable obstacle to existing systems. In addition, the dynamic nature of group interactions exacerbates the limitation, as the collective actions of individuals contribute to a complex amalgamation of signals that traditional systems find challenging to deconstruct.

Robust multi-user activity recognition systems will contribute to early fall detection in daily activities, indicating health issues or emergencies. Such applications underscore the potential impact of HAR on improving the quality and efficiency of healthcare services. Consequently, researchers delve into developing algorithms, drawing inspiration from machine learning techniques such as Support Vector Machine SVM [25], [26] and harnessing the power of deep learning models like Convolutional Neural Network CNN and recurrent neural networkRNN. Researchers are also looking into how to make personalised and adaptive recognition systems for each user, considering their unique movement patterns and behaviours [27], [28].

Depatla and Mostofi [24] introduced an innovative methodology for quantifying the number of individuals within a building by leveraging WiFi Received Signal Strength Indicator (RSSI) measurements and inter-event times. The proposed approach exploits the resilience of inter-event times to signal attenuation through walls, enabling the estimation of total occupants. The results reveal high accuracy in estimating the total number of people behind walls while minimizing the necessity for prior calibrations [29]. However, the current framework only considers scenarios with gradual changes in the total number of people over time, leaving out rapid fluctuations within the estimation period. Furthermore, Jing He and Yang developed a WiFi-based multi-user action recognition system called IMar [30]. The method involves building an amplitude relation model of a multi-person action scene and combining it with tensor decomposition to obtain continuous action data for each person. Using tensor completion makes the decomposed single-person data more informative, which is convenient for action recognition and counting and improves accuracy. The novelty of IMar lies in its ability to analyse and record the different actions of multiple users in a device-free scene. The study's limitations include the need for a clear line of sight between the WiFi access point and the users and the potential for interference from other WiFi signals.

Jianing Yu proposed a novel framework approach for WiFi sensing using 5G that fuses features from different bands and granularity levels to achieve superior performance in various sensing tasks [31]. The proposed method matches the feature granularity between WiFi channel measurements at 5 GHz and 60 GHz via a learning-based fusion block. The novelties of this method include achieving around 5% gain in accuracy on average over the best baseline methods available and mitigating the requirement of a large amount of training data. The method pre-trains a multi-band fusion network in an unsupervised fashion, fine-tuning each sensing head with limited labels. The need for additional validation in complex environments is due to further restrictions on the technique, which leads to low-performance classification.

Similarly, in their work, Ashleibta et al. [11] utilised a 5G-enabled RF-sensing system that works at 3.75 GHz and uses CSI signals to detect activity from multiple users simultaneously without touching the sensors. The proposed model uses a frequency range of 3.4 GHz to 3.8 GHz. The 5G bandwidth distinguishes it and represents a pioneering implementation of 5G sensing. The Ultra-wideband UWB method uses CSI signals to group activities into groups. It does this by combining 5G RF-sensing and deep learning models. However, the system requires a stable 5G signal, posing limitations in specific environments and experiencing reduced accuracy with an escalating number of users.

A novel approach for multi-person respiration sensing using commodity WiFi devices called MultiSense was developed by Zeng et al. [20]. They utilise Blind Source Separation (BSS) and ICA to extract respiration patterns from WiFi CSI signals. The authors demonstrate that MultiSense outperforms existing approaches regarding accuracy and scalability. However, the method faces limitations, including the need for prior knowledge of the number of people and its susceptibility to environmental factors like motion and interference. Additional work by Tan et al. [14] establishes an innovative system for tracking multiple users and recognising their activities using standard WiFi signals, addressing challenges posed by signal interference and noise. The methodology employs signal-processing techniques and machine-learning algorithms for feature extraction. However, the study includes experiments conducted in a controlled environment, raising concerns about the system's applicability in more complex real-world scenarios. Additionally, the system's accuracy fails to classify if a location is crowded with users.

Venkatnarayan et al. proposed a WiFi-based multi-user gesture recognition model called WiMU [9], emphasising the capability of signals to discern gestures simultaneously by multiple users. The researchers employed a blend of signal processing techniques and machine learning algorithms for precise gesture recognition across diverse user positions and orientations. Nevertheless, the method exhibits dependence on a stable WiFi signal, posing a potential limitation in less favourable environments. Furthermore, its accuracy declines with concurrently executed gestures, reflecting a common challenge in multi-user gesture recognition systems. Furthermore, [22] proposed a system that capitalises on the insight that WiFi signals capture the movements of all users, enabling the recognition of individual activities. This work details the extraction of features from WiFi signals to achieve accurate activity recognition for three users, with a commendable accuracy exceeding 90%, employing a combination of signal processing techniques and machine learning algorithms. However, the system faces potential performance degradation in environments characterised by high levels of interference or noise in WiFi signals. [22].

The real-world implications of inaccuracies in activity recognition [32] emphasise the pressing need for robust and scalable solutions in multi-user HAR. In scenarios where concurrent activities of multiple users unfold in real-time, misattribution or misinterpretation of actions leads to a cascade of errors, particularly compromising the overall utility and reliability of intelligent systems [33]. Failure to identify and distinguish these activities not only impedes the potential for personalised automation but also results in suboptimal user experiences. Furthermore, the increasing integration of multi-user HAR in domains such as healthcare and security underscores the importance of addressing this challenge. Similarly, recognising diverse individuals and their activities in security applications is pivotal for threat detection and anomaly identification. As a result, improving HAR systems' discriminatory capabilities in multi-user scenarios becomes a technological necessity for the practical application and implementation of smart environments.

The basis of WiFi sensing lies in analysing CSI signals, which encapsulate information about the channel's behaviour in the time and frequency domains. The CSI signal is represented as a matrix, typically denoted as $H(t, f)$, where t represents time and f denotes frequency. The variations in this matrix over time correspond to changes in the wireless channel due to the presence and movement of people. Detecting multiple individuals introduces a complex interplay of signals, leading to superimposed CSI patterns. This complexity amplifies the difficulty in isolating and characterising each individual's contribution to the overall signal, thus affecting the reliability of activity sensing, as shown in Fig. 2. The impact of multiple individuals on WiFi-based sensing is twofold. The superposition of signals necessitates addressing the equation $\mathrm{CSI}(t, f) = \sum_{i=1}^{N} H_i(t, f)$, wherein distinguishing the unique contribution of each individual becomes intricate and the mutual interference between individuals. Fig. 2 illustrates the reflected signals in a WiFi sensing region influenced by the presence of multiple individuals, represented by varying colours such as red, green, and blue. It emphasises the scattering effects of signals, which fluctuate due to the dynamic nature of human presence.

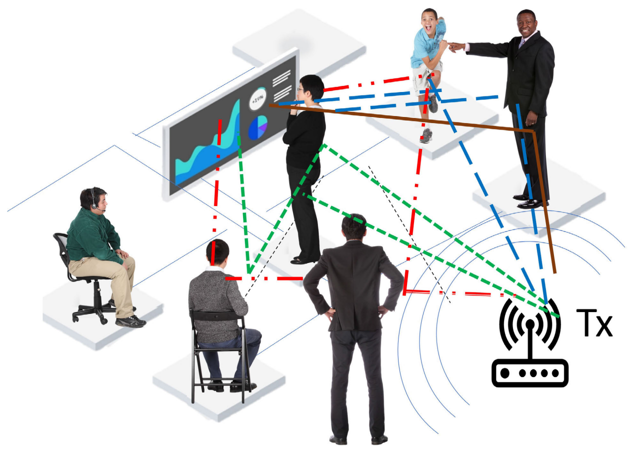

FIGURE 2. The composite Channel State Information signals arise as the number of individuals increase.

In the context of multi-user localisation, [34] introduces a novel scheme that utilises multi-radar cooperative sensing to enable continuous tracking within indoor settings. The article addresses the inherent difficulties in tracking multiple individuals by implementing a dual segmentation technique. The significance of this technology extends to diverse domains, such as security surveillance and location-based services. However, the proposed approach relies on complex systems utilising mmWave radar, which operates in the 60–64 GHz frequency band, rather than leveraging readily available WiFi infrastructure. Another model by Peng et al. [35] proposes an introduction gesture scheme to estimate the spatial channel of dynamic reflections under the impact of solid phase noise. They used the spatial beamforming technique to recognise the hand gestures of multiple users. A limitation of beamforming for multi-user sensing is the interference of signals and overlapping beams. That restricts individuals' accurate differentiation and weakens the sensing system's effectiveness.

Furthermore, in the context of multi-user localisation, Li et al. [34] employs a novel scheme using multi-radar cooperative sensing to enable continuous tracking within indoor settings. Their work addresses the inherent difficulties in tracking multiple individuals by implementing a dual segmentation technique. However, the proposed approach relies on complex systems utilising mmWave radar, which operates in the 60–64 GHz frequency band, rather than leveraging readily available WiFi infrastructure. To find the spatial channel of dynamic reflections when solid-phase noise is present, another study by [35] suggests a preamble gesture scheme. This scheme uses the spatial beamforming method to track the hand movements of multiple users. One limitation of employing beamforming for multi-user sensing is the potential for signal interference and overlapping beams. These factors impede accurate differentiation and tracking of individuals, thereby compromising the sensing system's effectiveness and reliability.

III. METHOD IMPLEMENTATION

A. OPERATIONAL WORKFLOW

The algorithmic system architecture begins with the data collection module, which acquires diverse datasets containing relevant information for the task. Following this, a meticulous pre-processing stage cleanses and organises the raw data, ensuring its suitability for subsequent processing. The next feature extraction phase employs sophisticated techniques, including ICA and CWT, to purify meaningful patterns and representative features from the pre-processed data. The algorithm establishes spatial correlations and similarities between the extracted features at the matching locations. The algorithm incorporates both ICA and CWT features to enhance the robustness and richness of the representation. The pivotal component of the architecture is an LSTM network. This deep learning model is adept at capturing temporal dependencies and nuances in the data. The mechanism further refines the model's focus on salient features, optimising its ability to discern complex patterns. The LSTM classifier predicts using the learned representations, offering a comprehensive and robust solution for the specified task. Fig. 3 shows the algorithmic system architecture, which looks like a structured flowchart. It has blocks that show data collection, pre-processing, feature extraction using ICA and CWT, matching locations, and an LSTM classifier that offers a complete and sequential method for strong pattern recognition.

B. DATA COLLECTION

In wireless communication systems, CSI is a metric that characterises the dynamic nature of the communication channel between a transmitter and receiver. The mathematical representation of CSI encapsulates the underlying in (1), where Y denotes the received signal, X represents the transmitted signal, H signifies the channel matrix reflecting the CSI, and N means the inherent noise in the channel [36].

$$Y = HX + N \tag{1}$$

The CSI equation encapsulates the intricate interplay of signals, making it a foundational concept in wireless signal processing. Collecting CSI data involves meticulous measurements of the received signal under varying conditions, yielding a dataset $\{Y1, Y2, \ldots Yn\}$ [37]. Each Yi represents a vector capturing the received signal at a specific instance. CSI data collection involves sending known injected data and monitoring system context to extract CSI, such as Nexmon, and then observing the received signal [38], [39]. CSI data extraction algorithms utilise the collected data to form the channel matrix Hi in (2). This dataset often represented as a three-dimensional matrix across subcarriers and time snapshots, is integrated to understand and optimise wireless communication systems, forming the basis for further analysis and enhancement.

$$H_i = \begin{bmatrix} h_i^{11} & h_i^{12} & \ldots & h_i^{1N_T} \\ h_i^{21} & h_i^{22} & \ldots & h_i^{2N_T} \\ \vdots & \vdots & \vdots & \vdots \\ h_i^{N_R1} & h_i^{N_R2} & \ldots & h_i^{N_RN_T} \end{bmatrix} \tag{2}$$

The process of collecting CSI entails the capture of wireless signals, with a particular emphasis on an 80 MHz bandwidth. The mathematical part of this process involves using signal processing transforms, such as Fourier transforms, on the raw signal to separate the important CSI information from the wider WiFi signal. Researchers often employ tools such as Nexmon [38], OpenWRT [40], and Atheros tools [41] to facilitate the extraction and analysis of CSI data. We accomplish the separation by exploiting distinctive CSI characteristics, such as its unique frequency spectrum and temporal variations [42]. After successfully isolating the CSI data, we graphically represent it by plotting amplitude against time or frequency. The graphical depiction offers a visual insight into the dynamic changes in CSI amplitude, providing valuable information about the wireless channel's behaviour [30], [43].

C. PREPROCESSING

Implementing data preprocessing improves the accuracy and dependability of the amalgamated CSI data. The procedure must separate CSI data from nearby WiFi router signals to extract relevant information. Applying signal processing techniques and filtering algorithms discriminates against non-CSI components, achieving the task. Initially, a median filter strategy eliminated potential aberrations in the data by substituting anomalous values with the median derived from adjacent data points. Despite the initial filtering, some outliers persisted, prompting Hampel to fill outlier filters to mitigate their impact by replacing them with the original non-outlier values in the dataset. Hampel used a moving median filter technique to refine the data and diminish

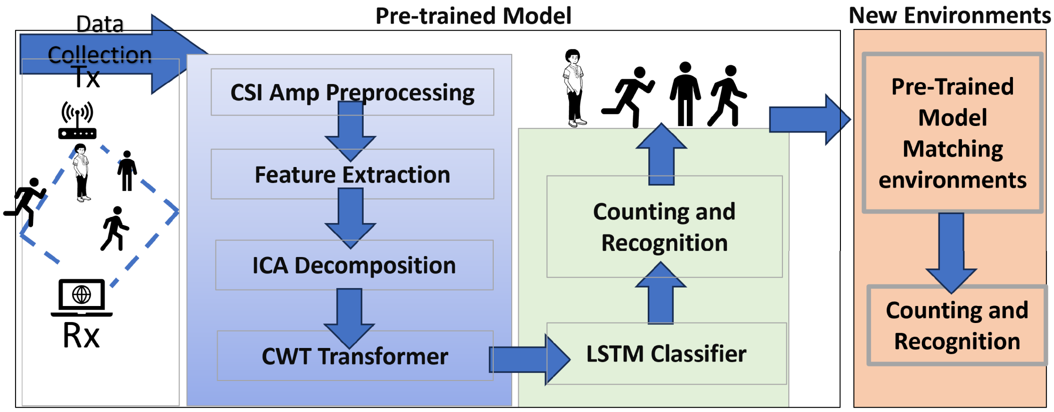

FIGURE 3. The flowchart for multi-user sensing includes the pre-trained model, feature extraction of ICA and CWT features.

residual noise by computing the median value across data points [36]. The applied filtering approach mitigated minor fluctuations or inconsistencies in the data. The median filter is a non-linear technique for suppressing impulsive noise or outliers within a signal. The mathematical representation of the median filter is shown in (3).

$$y[n] = \text{median}(x[n-m], x[n-m+1], \ldots, x[n+m]) \quad (3)$$

Here, $x[n]$ represents the input signal, and m is the filter window size. The Interpolation step is introduced to handle outliers in the data, replacing them with interpolated values based on neighbouring points. The interpolation is represented mathematically in (4).

$$y_{\text{interp}}[n] = y[n-1] + y[n+1] \cdot 2 \quad (4)$$

Equation (3) computes an interpolated value y_{interp} $[n]$ as the average of the adjacent data points $y[n-1]$ and $y[n+1]$. This technique used to smooth the CSI data, mitigating occasional aberrations and improving the overall quality of the signal. Smoothing involves reducing high-frequency noise and emphasizing underlying trends in the data. Smoothing is a common preprocessing step in signal analysis to highlight relevant patterns while suppressing noise. The moving average filter enhances the smoothness of the CSI data further, and The formulation of the moving average filter is represented by (5):

$$y_{\text{avg}}[n] = \frac{1}{N} \sum_{k=0}^{N-1} x[n-k] \quad (5)$$

N represents the filter length, and $x[n]$ is the input signal. In contrast, the moving average filter operates by calculating the average of the current and $N-1$ previous data points. When used with these filters, they play a crucial role in reducing noise and refining the CSI data for subsequent analysis, as shown in Fig. 4. The figure comprises three subplots illustrating a signal's preprocessing and filtering stages. The subplot in Fig. 4(a) demonstrates the removal of null and pilot subcarriers; Fig. 4(b) shows the signal with noise before filtering; and Fig. 4(c) represents the filtered CSI amplitude signal.

D. ICA AND CWT

ICA is a statistical signal processing technique that seeks to uncover independent sources from observed mixed signals. Considering the matrix H captures the linear relationship between the observed mixed signals, represented by the matrix X, and the separate source signals, represented by the matrix S, given in (6):

$$X = HS \quad (6)$$

S is an $M \cdot T$ matrix of mixed signals, M is the number of antennas, T is the number of samples, and S is an $M \cdot T$ matrix of independent source signals. The objective of ICA for CSI is to estimate the demixing matrix W such that $\widetilde{S} = W \cdot H$. The columns of $\widetilde{S}$ represent the estimated independent source signals. The ICA algorithm aims to maximise the statistical independence of the estimated sources. In this work, we implemented diverse strategies and techniques to address our research's specific challenges and objectives. The FastICA algorithm is a computationally efficient approach to solving the ICA problem, and its application to CSI data involves specific adaptations [20]. Whitening, a foundational step, transforms the observed CSI matrix H into uncorrelated signals with unit variance, facilitating subsequent processing. Negentropy maximisation is the focal point of FastICA, aiming to enhance the non-Gaussianity of the whitened CSI matrix.

The whitening process for CSI matrices involves transforming the observed CSI matrix H into uncorrelated signals with unit variance. The covariance matrix E of H computes the whitening process and obtain the whitening matrix $\sqrt{E}$. The expression for the whitened CSI matrix $\widetilde{H}$ is described in (7).

$$\widetilde{H} = H \cdot \sqrt{E} \quad (7)$$

The whitening step is pivotal in enhancing the subsequent processing of the CSI data. It ensures that the transformed

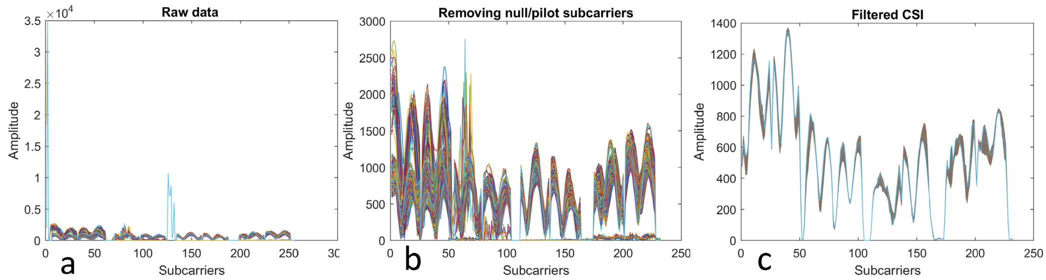

FIGURE 4. Filtering and preprocessing of the CSI amplitude signal through denoising and outlier removal.

signals are decorrelated and possess a standardized variance. Furthermore, the Negentropy maximization is the core objective of the FastICA algorithm when applied to CSI data. It seeks to maximize the non-Gaussianity of the whitened CSI matrix $\widetilde{H}$. The optimization problem represents the introduction of a non-linear function g to approximate negentropy in (8).

$$W^* = \text{argmax}_W \text{ Negentropy}(\tilde{H}) \tag{8}$$

The iterative nature of the FastICA algorithm involves updating the unmixing matrix W using a gradient ascent approach is represented by (9).

$$W \leftarrow W + \alpha \left(E\{H\tilde{g}(W^T\tilde{H})\} - E\{g'(W^T\tilde{H})\}W \right) \tag{9}$$

where α is a step size parameter, and g' denotes the derivative of the non-linear function g. The process continues until convergence, resulting in an optimised unmixing matrix W that maximises the non-Gaussianity of the whitened CSI signals. Negentropy maximisation reveals the independent components within the CSI data, extracting meaningful information from complex composite signals, and separating them purely.

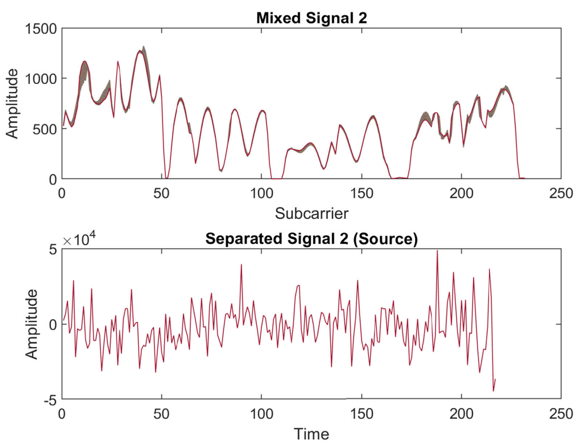

FIGURE 5. Plot of the separated signal using ICA in CSI data.

The process involves optimizing a non-linear function approximating negentropy, which is pivotal for revealing the independent components in the channel. The algorithm improves an unmixing matrix W using gradient ascent. It optimizes the negentropy of the whitened CSI matrix, as depicted in Fig. 5. Normalization ensures that the unmixing matrix maintains orthogonality using (10).

$$W \leftarrow \frac{W}{\|W\|} \tag{10}$$

Subsequently, we combine the unmixed signals into a matrix, denoted as both, and apply the Continuous Wavelet Transform (CWT) to this matrix. The CWT is a mathematical operation that decomposes a signal into its constituent wavelet components. In this context, it allows for a more detailed analysis of the frequency content of the unmixed signals. The CWT is a powerful tool that enables simultaneous analysis of signals in both time and frequency domains, as illustrated by (11).

$$CWT(a, b) = \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-b}{a} \right) dt \tag{11}$$

where $x(t)$ represents the signals, $\psi(t)$ is the analysis, a is the scale parameter, and b is the translation parameter. The integral evaluates the possible values of a and b representing the signal in the time-frequency domain. The scale parameter determines the width of the wavelet, which influences the balance between frequency and time localization. A smaller one offers higher frequency resolution but inferior time localization, and vice versa.

E. FEATURE EXTRACTION

LSTM represents a recurrent neural network distinguished for its adeptness in capturing extended dependencies within sequential data. In WiFi-based sensing, LSTM serves as a potent tool for feature extraction. It leverages attributes like memory retention, non-linear mapping, and the ability to sequence. The LSTM cell's structure also enables it to remember pertinent information, discern intricate relationships and temporal dynamics, and accommodate diverse sequences. Unlike conventional techniques and simpler models, LSTM excels in discerning patterns, deriving meaningful representations, and exploiting temporal dependencies. Therefore, it is well-suited for feature extraction in CSI analysis. To preserve the integrity of the acquired data and forestall the loss of pivotal signal features, we consciously abstain from using dimensionality reduction techniques such as PCA or Linear Discriminant Analysis (LDA). Instead, we opted to employ the complete dataset by organising and correlating it based on the amplitude of the CSI.

F. LSTM CLASSIFIER

The proposed architectural design utilises LSTM layers in a sequential framework for classification, as illustrated in Fig. 6. Memory cell integration into the LSTM layer of this framework enables effective information retention and processing over time. The capability is particularly crucial in scenarios where temporal relationships play a paramount role, permitting the model to capture and model long-term dependencies in the data. A dropout layer follows the LSTM layer to address the potential overfitting issue and enhance generalisation. The next layer, the ''Fully Connected LSTM Layer,'' performs fully connected operations on the output from the preceding LSTM layer. This process facilitates the discernment of intricate relationships and complex patterns within the data. The model then incorporates an output layer comprising two fully connected neurons, serving as a classifier. Lastly, a fully connected layer precedes a SoftMax layer, generating probability distributions across the seven predefined output classes. The behaviour recognition system employs LSTM networks to extract features from signal data. The presented architecture outlines a sequential model that employs LSTM layers for classification, as illustrated in Fig. 6. The first layer receives data input with the sequence's length to initiate the workflow. Subsequently, this sequential data is processed by an LSTM layer, a specialised recurrent neural network layer adept at handling sequential information. Notably, the LSTM layer incorporates memory cells to retain information over time, capturing long-term dependencies within the data. After the LSTM layer, we include a dropout layer to introduce regularisation and address the issue of overfitting. Dropout selectively deactivates specific input units during training, thereby fostering the independence of neuron learning.

During the subsequent phase of training the model and fine-tuning hyperparameters, we partition the original training dataset into 80% for training and 20% for validation to assess the model's performance. The LSTM model was evaluated using the validation dataset, and the hyperparameters of the trained models were refined using an optimization approach. Additionally, we compare the performance of the hyperparameter-tuned models by evaluating them against the test results in the context of activity recognition.

G. MATCHING ENVIRONMENTS

To represent the variance between signals in two different environments, we use the mathematical form in (12).

$$\bar{v}_i = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu) \tag{12}$$

where $\bar{v}_i$ is the variance of N numbers of samples between the CSI amplitude at the first location x_i and the mean samples of the new location μ. The averaging process in (11) to measure the variation of signals, which provides features of attenuation variance spread between amplitudes of each subcarrier due to changes in surroundings. To express the process of matching signals by incorporating the calculated variance $\bar{v}_i$ with the new environment, we utilise (13) as follows:

$$Q_{\text{new}} = (\bar{v}_i) + \frac{1}{N}\sum_{i=1}^{N} Nv \tag{13}$$

The matching of the variance between the new environment signal, Nv, and the variance between the new and old signal $\bar{v}_i$ of the pre-trained dataset is shown by Q_{new} in (12). Here, N is the number of samples. We match the signals between the two environments by adding the variance from the pre-trained dataset to the new environment signal. The approach assumes a normal distribution for the signals and knowledge of the mean CSI heatmap of the empty orientation of the new environment. The pseudocode algorithm for CSI-based HAR involves capturing CSI data, preprocessing the data to remove noise and interference, extracting relevant features, applying a machine learning model to classify the activity, and outputting the recognized human activity as shown in Algorithm 1.

Step Algorithm 1: Multi-users WiFi HAR

Input: Raw CSI data $\leftarrow$ local median CSI $\overline{\text{Amp}}^1$
Output: Multi-users Activity classifications location independently
1: CSIAmp Hi $\leftarrow$ CSIraw - CSI pilot - CSI null to remove pilot and null subcarriers
2: Compare the current *CSIAmp* (i) with σ_{i} using Eq. 6
3: **if** *CSIAmp* $\text{k}_{\text{k}}^{\text{i}} - \mathit{CSIAmp}^{\ 1} > \text{k}\sigma \times \sigma_{\text{i}}$
4: $\text{CSI}\overline{\mathit{Amp}}^1 = \mathit{CSIAmp}(i)$
5: **end if**
6: $\text{CSI}_{\text{Amp}}(\text{i}) \leftarrow \text{CSI}\,\overline{\text{Amp}}^1$ outlier Hampel filter
7: $\text{CSI}_{\text{Amp\_hampel\_Denoised}}(\text{i}) \leftarrow \mathit{CSIAmp}(\text{i})$ median filter
8: Scout $[242 \times 500] \leftarrow$ Feature extraction from CSI Amplitude
9: Separate CSI signals using ICA $W \leftarrow W$ Eq. 8
10: Apply wavelet $W \leftarrow \frac{W}{\|w\|}$ to represent the separated data
11: Utilise LSTM model to extract 3D-CWT features for the Classification
12: Match environments by finding the variance $\bar{v}_\imath$ between two locations Eq. 11

IV. IMPLEMENTATION AND EVALUATION

A. EXPERIMENT SETTINGS

For implementation and evaluation, we explored the model's design, training methodology, and rigorous performance evaluation against established benchmarks. The experiment included scenarios with varying numbers of participants to enhance the diversity of the dataset, different activity intensities, and overlapping actions. The participants moved freely within the designated space while the WiFi sensors continuously recorded the variations in signal patterns. The experiment uses self-collected datasets that addressed the scarcity of relevant datasets in the field of multi-person human activity recognition, and tested them in real-world settings. In the context of multi-person human activity recognition through WiFi signals, we set the experimental configurations methodically to moderate the absence of

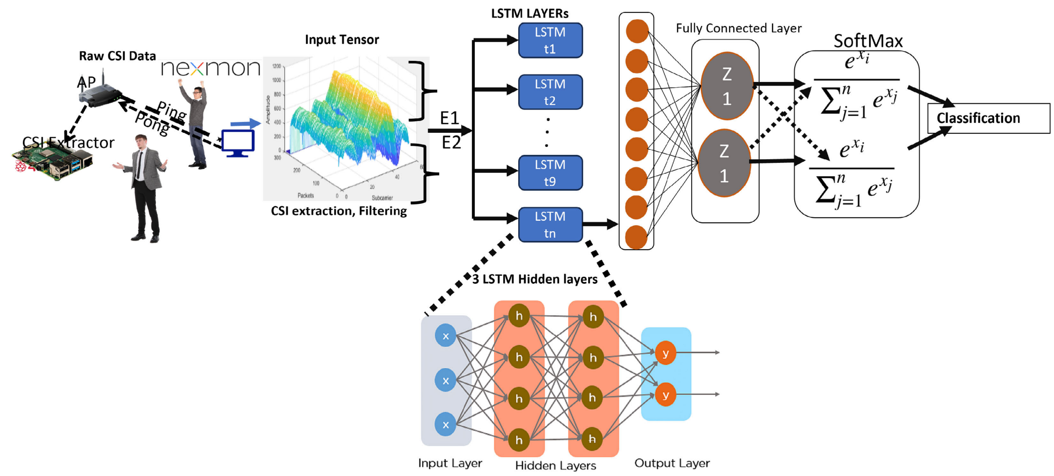

FIGURE 6. Illustration of the applied architectural layers of a Long Short-Term Memory (LSTM) network.

shared datasets tailored for this specific domain. Given the acknowledged challenges of acquiring authentic data encompassing diverse human interactions, the work adopted a self-collected dataset. The data collection entailed strategically deploying a network of WiFi sensors throughout the experimental environment, ensuring comprehensive coverage and consistent signal strength across the spatial expanse. We configured the CSI sensors to capture fluctuations in WiFi signals attributable to human activity, yielding a dataset that encapsulated a spectrum of activities.

The assessment included examining data acquired through the Nexmon CSI extraction tool. Nexmon facilitated the retrieval of CSI data from a Raspberry Pi 4B explicitly configured for very high-throughput mode, with a bandwidth of 80 MHz. The obtained CSI samples provided detailed channel information, including 256 subcarriers. The data collection utilised clustered Raspberry Pi units to enable two receivers within the RPI system, as shown in Fig. 7. The experimental arrangement featured a Broadcom BCM43455c0 NIC with a Raspberry Pi 4B unit as the receiver and a TP-Link AC1350 router as the transmitter. Both Raspberry Pi devices operated on Linux version 5.10.92 firmware. The receiver and transmitter supported multi-user MIMO functionality. The configuration of the experimental setup ensured compliance with industry standards, thereby augmenting the reliability and validity of the evaluative procedures. The CSI datasets we got were from different indoor settings, specifically inside a residential apartment hall, which is limited by the spatial parameters shown in Fig. 7.

The Nexmon firmware is used for monitoring the router and extracting CSI data from TCPDUMP files, and processing it with MATLAB scripts for rigorous cleaning and preprocessing. It covers a variety of scenarios and activities in controlled environments with specific room dimensions (9 m x 13 m x 30 m) and device placements (3 m distance between Tx and Rx, 1 m height for transmitters, 1.5 m for receivers) as shown in 7. Distinct labels categorise

FIGURE 7. Experiment layout: (a) real location, (b) 2D layout, (c) 3D simulation.

each scenario, indicating various conditions such as empty location, standing activity, walking, sitting alone or with others, fall, and running. Additionally, the dataset includes mixed activities involving walking, standing, sitting, falling and running performed by one or more individuals.

The dataset consists of CSI measurements collected by four individuals across various scenarios to capture diverse human activities and environmental interactions. The activities encompass a spectrum of motion states, including stationary, empty, standing, sitting, walking, running, and combinations thereof, thereby ensuring a comprehensive representation of human interactions within the dataset. Table 1 summarizes the dataset collection conducted using WiFi devices equipped with multiple antennas to ensure comprehensive CSI data with following annotation (E = Empty; ST = Stand; W = Walk; R = Run; Si = Sit;F = Fall). This setup enabled the capture of fine-grained variations in the wireless signals, reflecting the distinct movement patterns and interactions in each scenario. The resulting dataset provides a rich resource for analysing human activities and developing models for activity recognition and environmental sensing using WiFi signals. The datasets used in this study are available at this *linked*[1] to facilitate further investigation, replication, and potential collaborations.

B. MULTI-USER PERFORMANCE EVALUATION

The experimental framework suggested using a combination of ICA and CWT analyses to identify human activity within the context of WiFi signal analysis. The primary objective was to enhance the system's discriminatory capability in discerning various human activities. We employed ICA statistically to isolate independent sources within the observed WiFi signals. ICA's methodological choice aimed to reduce signal complexity and enhance the system's ability to differentiate distinct human activities. The analysis facilitated a more refined understanding of the underlying signals by extracting statistically independent sources, thereby improving recognition accuracy. The challenge of distinguishing between three distinct scenarios, such as a single person walking or more people walking, arises from the intricate nature of mixed signals in human activity recognition. The complexity lies in the overlapping patterns within the signals emitted during these walks, as shown in Fig. 8.

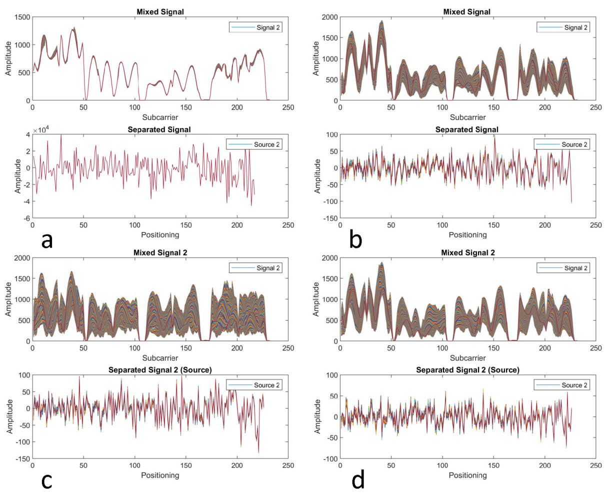

FIGURE 8. The subsequent ICA separation illustrates the complexity of disentangling the composite signals. Subplots show the unique signals corresponding to (a) standing activity, (b) single-person walk, (c) two-person walk, and (d) three people walking. The signals show how ICA provides better analysis of multi-user feature patterns.

Fig. 8 (a) depicts the signal separation through ICA after illustrating the composite nature of mixed signals in Fig. 8. Notably, Fig. 8 (b), (c), and (d) correspond to scenarios involving single-person, two-person, and three-person walking, respectively. Each method generates distinctive signals indicative of the respective walking configurations. However, the convergence of these signals introduces a notable level of intricacy that challenges the recognition process. The nuanced variations in gait patterns, stride lengths, and spatial dynamics across the distinct walking scenarios contribute to a complex signal amalgamation. This intricate interplay within the mixed signals underscores the difficulty of discerning and classifying the diverse walking configurations.

The combination of signals underlines the difficulty in unravelling the distinctive features that characterise each walking scenario. Addressing these challenges requires enhanced signal processing techniques and robust algorithms to tease apart the overlapping elements and improve the precision of recognition in dynamic and mixed-signal environments. Additionally, we applied CWT analysis to provide a time-frequency representation of the WiFi signals. The time-frequency presentation provided insights into the temporal dynamics of human activities, resulting in a more comprehensive and detailed understanding of the material aspects involved in the recognition process. Using both ICA and CWT analyses at the same time revealed a complex analytical framework and a way to recognise human activity through WiFi signal analysis that works well together. The combination of methods led to a better understanding of the complex time patterns that make up different human activities in dynamic environments.

The experiment investigates the system's ability to identify and classify an individual's activities within the monitored environment. Furthermore, the technique served as a foundational assessment to establish the baseline performance of the activity recognition system. The CWT indicates how static activity, single-person walking, two-person walking, and three-person walking have different time-frequency and magnitude characteristics, as seen in Fig. 9. The subplots show the movements of human activities over time by using the wavelet to break the signals down into different frequency components. Wavelet analysis helps identify the unique patterns and traits associated with each activity. It enables a more accurate and detailed analysis of the temporal dynamics and variations in human movement.

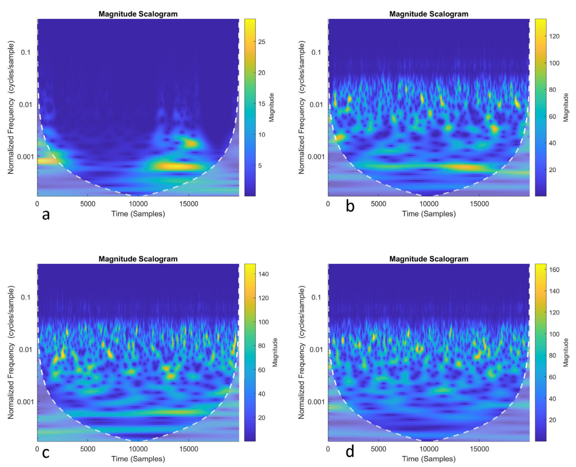

FIGURE 9. CWT representation for various activities: (a) static activity, (b) single-person walk, (c) two-person walk, and (d) three-person walk. The plots capture unique features distinguishing specific human activities.

[1] https://data.mendeley.com/preview/bkgw7c57wf?a=7d5f919f-a75e-43cd-89c6-ef596662477d

TABLE 1. Dataset table for multi-user scenarios.

scenario / No. People	1	2	3	4	5	6	7	8	9	10	11	12	11	12	13	14	15	16	17
1	E	ST	ST	ST	W	W	W	Si	Si	Si	R	R	R	R	W	W	R	W	R
2	E	E	ST	ST	E	W	W	E	Si	Si	E	E	R	R	ST	W	W	F	R
3	E	E	E	ST	E	E	W	E	E	Si	E	E	E	R	Si	ST	ST	S	W
Packet injection	$500 packets/$s																		
No. antennas	2																		
Freq	5GHz																		
BW	80MHz																		

Furthermore, the implementation of 3D CWT represents a notable improvement in the signal processing stage, particularly in enhancing the discrimination between complex signals. Unlike traditional 2D approaches, the 3D CWT extends the analysis into both time and frequency dimensions, providing a more comprehensive understanding of signal characteristics. By capturing the dynamics of movements across both time and frequency axes, the 3D CWT makes it easier to tell the difference between signals more nuancedly. This leads to higher precision in signal processing applications, especially when there are complex changes in both time and frequency. Fig. 10 shows the capability of 3D-CWT in sensing static and dynamic activities such as standing and walking. The implementation of ICA with 3D-CWT shows considerable differences in signals related to these activities for different numbers of people, which enables better signal sensing for multi-users.

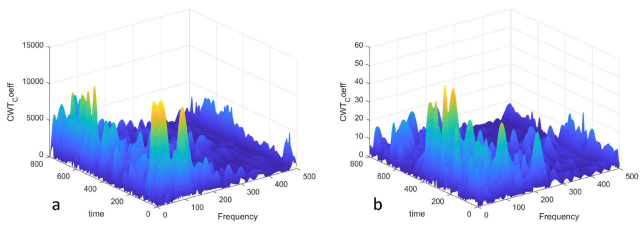

FIGURE 10. The 3D CWT shows dissimilarities and variances between (a) static and (b) walking activities.

The ICA has also proven to be a robust method for segregating signals derived from diverse human activities, particularly distinguishing between static and dynamic activities. In static activities, ICA separates signals associated with stationary postures, isolating the intrinsic features indicative of minimal movement. In contrast, ICA separates signals characterised by dynamic changes and varying motion patterns in dynamic activities. This highlights such activities' inherent temporal and spatial complexities, as depicted in Fig. 11.

Fussing ICA and CWT delivers a sophisticated approach to disentangling and analysing CSI signals to discern the number of individuals engaged in activities. Fig. 12 demonstrates the effectiveness of this implementation, showcasing the distinct contributions of ICA and CWT. Notably, the comprehensive integration of ICA and CWT emphasises their complementary roles in unravelling the intricacies of CSI

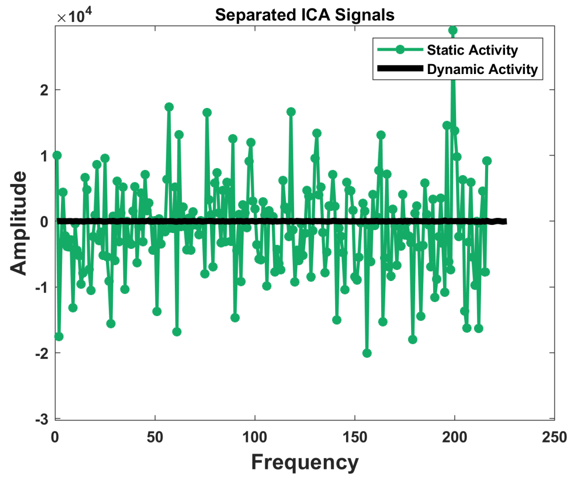

FIGURE 11. The ICA separation of signals illustrates the dissimilarity between static and dynamic activities. The way ICA is used in methodology separates signals, showing the features of still positions in static activities and capturing the changing patterns and changes that happen in dynamic activities.

signals, ultimately contributing to a more precise and accurate assessment of the number of people involved in the observed activities. The integrated methodology enhances the system's capability to distinguish the number of people engaged in activities, demonstrating the complementary roles of ICA and CWT in signal analysis.

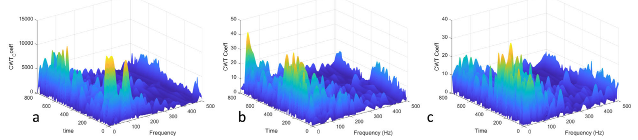

FIGURE 12. The unique analysis of the number of people walking through the combined implementation of ICA and CWT shows the combined approach. First, ICA breaks down CSI signals into statistically independent components. Then, CWT coefficient analysis allows for a more detailed evaluation of changes in time and frequency. (a) represents one person standing, (b) one person walking, and (c) two people walking.

C. IMPACT OF MULTI-PEOPLE MOVEMENT

The evaluation results indicated that the proposed method impacted the accurate classification of groups' movement activities. In the single-user scenario, the system recognised and classified individual activities. Transitioning is more complex for two- or more-person scenarios; the process exhibited resilience in distinguishing overlapping signals and varying trajectories, showcasing its ability to classify group movements. The system's robustness in handling concurrent activities by multiple individuals was positive outcome, affirming its practical utility in scenarios involving collective actions. The experimental findings also highlighted the method's ability to count the number of people engaged in group activities. Functionality holds paramount importance in applications such as crowd monitoring and management. The evaluation demonstrates that the method addresses the challenges of classifying group movement activities. The method's success reflects its potential for practical scenarios requiring accurate and dependable recognition of human activities, even in the presence of multiple individuals. The proposed model excels at capturing dynamic alterations within the signal domain, especially in scenarios involving varying numbers of individuals engaged in activities such as walking. As individuals move, the CSI signal's dynamic changes exhibit discernible frequency patterns over time. Notably, the frequency changes are more pronounced when multiple individuals are involved in dynamic activities, such as walking. As a result, the ICA detects the higher frequency components associated with an increasing number of people participating in such activities. The ICA leverages dynamic frequency fluctuations over time, making it a powerful tool for differentiating the complexities that arise from varying numbers of users. As a result, this improves the CSI signal analysis and distinguishes between various individuals in different situations. Fig. 13 shows the separated walking signals performed by individuals, pairs, and groups of three, highlighting noticeable differences in the CSI signal plots. As more individuals participate in walking activities, the frequency of dynamic changes increases clearly, demonstrating the analysis's effectiveness in accurately collecting and differentiating between various scenarios.

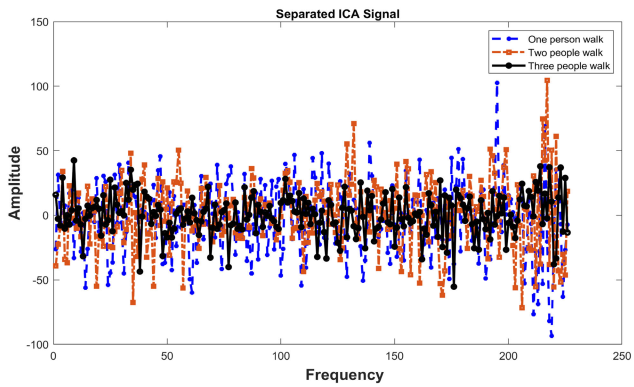

FIGURE 13. **Separated Signal variations in frequency change over time, highlighting the ability of ICA to distinguish between different numbers of individuals engaged in walking activities.**

The evaluation method focused on the system's ability to determine the number of individuals present within the monitored area. The results demonstrated the system's capability to count the number of people in diverse scenarios, including situations involving multiple individuals engaged in different activities. The results confirmed the system's effectiveness in discerning individual and collective human activities and underscored its practical utility in real-world settings.

D. COMPLEX ACTIVITIES AND ENVIRONMENTS EVALUATIONS

The proposed model addresses the intricate challenge of discerning between different and complex activities, particularly those that share common features, such as running and walking. The difficulty distinguishing between these activities is due to the similarity in specific motion patterns and feature representations. Moreover, integrating attention mechanisms enhances the model's ability to focus on critical features within the signal that are indicative of specific activities. Attention mechanisms enable the model to prioritise and emphasise unique characteristics by changing the weights of different signal components. The attention mechanisms act as a rational filter, supporting the model in selecting relevant features.

The evaluation is concerned with the ability to distinguish between complex activities and complex scenarios. It assesses the models' capacity to differentiate between intricate movements, providing insights into their robustness in addressing the intricacies of human behaviour within dynamic environments. The model demonstrates a remarkable ability to distinguish between complex activities shown in Fig. 14, such as walking and running, even when performed by different numbers of individuals. It announces the development of robust systems capable of nuanced human behaviour recognition in real-world, multi-faceted settings.

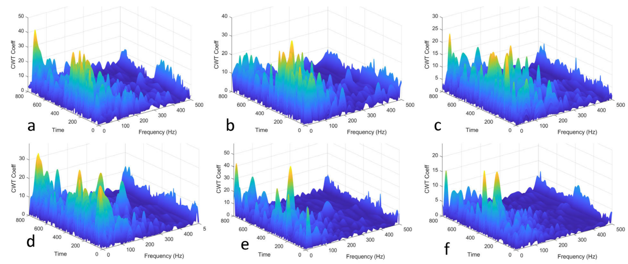

FIGURE 14. **Comparative analysis illustrating 3D-CWT classification of (a) one person, (b) two people, and (c) three people walking, followed by (d) one person, (e) two people, and (f) three people running.**

The discrimination between intricate activities such as walking and running, especially when involving varying numbers of individuals, poses a formidable challenge to activity recognition systems. The dynamic nature of these movements, encompassing distinct gait patterns, speeds, and accelerations, introduces complexity in signal analysis.

In complex environments, the integration of ICA with CWT provides advantages for signal processing, as illustrated in Fig. 15. We tested the model in three different locations with different dimensions: the classroom, and the home demonstrated in Fig. 16. The method separates the mixed signals into statistically independent components, which

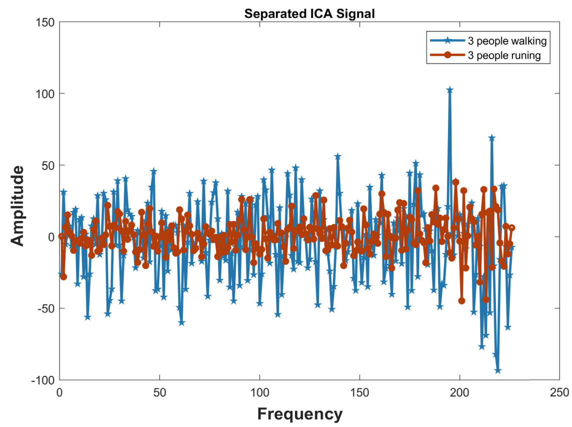

FIGURE 15. The plot shows the distinction between walking and running activities performed by three individuals. The robustness of the complicated human motion recognition methodology is evident in the classification of complex activities.

is important for isolating the desired signal from noise and interference. The synergy of ICA and CWT ensures robust performance in dynamic and cluttered environments, enabling precise and reliable signal analysis.

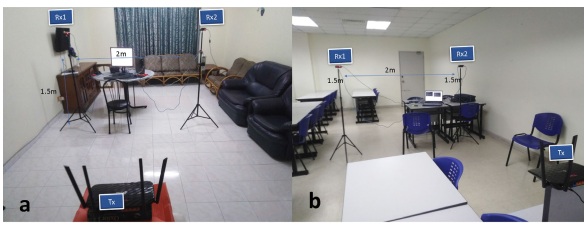

FIGURE 16. Different environments used for analysis: (a)Home, (b) Classroom.

The proposed approach enables for a precise characterization of the intricate dynamics inherent in such activities and in complex environments. It demonstrates the model's proficiency in navigating the complexities of diverse and simultaneous human movements. The plotted results in Fig. 15 depict the model's capability to distinguish between three people walking and three people running, showcasing its efficacy in handling complex activity recognition scenarios.

E. METRICS EVALUATION

In this work, the hyperparameters of the implemented models and training process are carefully selected to optimize performance and robustness. The SVM and Naive Bayes classifiers are evaluated with optimized configurations, employing confusion matrices for performance assessment. For the deep learning models (LSTM, GRU, Bi-LSTM, and Attention Mechanism), the architecture is defined with specific configurations: sequence input size of 50, varying numbers of hidden units (128, 100, and 128). During training, the Adam optimizer is utilised with a gradient threshold of 1. The models undergo training over 50 epochs with a minibatch size of 32, and an initial learning rate of 0.0001. Training progress is monitored and visualized using MATLAB's, while validation is performed on separate data partitions, ensuring model generalization.

The comparative analysis, including classifiers such as SVM, Naïve Bayes (NB), LSTM, Gated Recurrent Unit Networks GRU, Bidirectional LSTM (BiLSTM), and attention mechanisms, is pivotal for understanding the nuanced performance of each approach. These models excel at capturing temporal dependencies and contextual information crucial for accurate activity recognition in scenarios with multiple participants. Attention mechanisms improve the model's ability to focus on important data features and relationships, leading to more accurate multi-person sensing than regular classifiers. Classifier performance is evaluated in Fig. 17, comparing the accuracy achieved by SVM, Naive Bayes NB, LSTM, GRU, BiLSTM, and LSTM Attention Mechanism models. The analysis offers insights into each model's relative strengths and capabilities for classifying data.

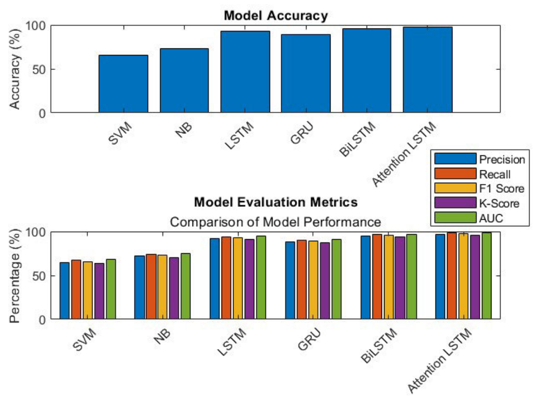

FIGURE 17. The classifier performance assessment compares the accuracy achieved by SVM, NB, LSTM, GRU, BiLSTM, and Attention Mechanism models.

The performance of algorithms was evaluated based on multiple metrics: accuracy, precision, recall, F1 score, K-score, and AUC. Precision (P) is defined as the ratio of true positive (TP) predictions to the total number of positive predictions:

$$P = \frac{TP}{TP + FP} \tag{14}$$

Recall (R), also known as sensitivity, is the ratio of true positive predictions to the actual number of positive instances:

$$R = \frac{TP}{TP + FN} \tag{15}$$

The NB model performed better, achieving an accuracy of 73%, with precision and recall of 72% and 74%, respectively. The F1 score for NB was 73%, where the F1 score ($F1$) is the harmonic mean of precision and recall:

$$F1 = 2 \cdot \frac{P \cdot R}{P + R} \tag{16}$$

Advanced deep learning models exhibited higher efficiency. The Long Short-Term Memory (LSTM) network achieved an accuracy of 93%, with precision and recall both around 92% and 94%. The Gated Recurrent Unit (GRU) model closely followed with an accuracy of 89%, precision of 88%, and recall of 90%.

The Bidirectional LSTM (BiLSTM) model further enhanced performance, achieving 95% and 97% precision and recall, respectively. The highest efficiency was exhibited by the Attention LSTM model, with an accuracy of 97%, precision of 97.5%, and recall of 98%. Furthermore, the K-score, another measure of the classifier's performance, was lower than precision and recall but followed similar trends. AUC, which measures the area under the ROC curve, confirmed the superior performance of deep learning models, especially the Attention LSTM, with an AUC of 98%.

Overall, while traditional machine learning models like SVM and NB showed moderate performance, the deep learning models, particularly BiLSTM and Attention LSTM, outperformed them across all metrics, demonstrating superior capability in complex environments and tasks. Moreover, these models often consider computational efficiency when there is an observed increase in accuracy. The variance in computational efficiency is primarily caused by the number of hidden layers and the overall complexity of the model architecture. While LSTM, GRU, BiLSTM, and Attention Mechanisms offer superior accuracy, their computational demands can be higher due to increased model intricacy. Furthermore, the observed trends in the box plot in Fig. 18 show a corresponding drop in accuracy when the number of activities or people increases. It indicates the increased difficulty of recognising diverse activities or handling multiple individuals. Conversely, as the number of samples increases, the accuracy tends to improve, emphasising the positive impact of larger datasets on model performance. The graphical representation, labelled axes, and title enhance the interpretability of how these factors influence the model's accuracy.

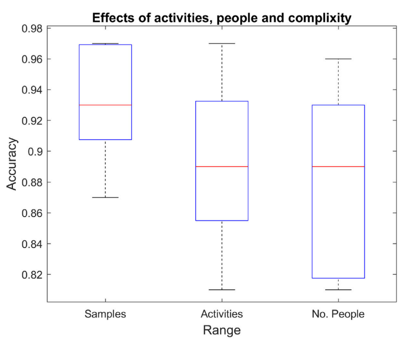

FIGURE 18. Model accuracy exhibits sensitivity to activity, people, and samples.

F. MODEL ADVANCEMENTS

The analysis of different indoor human activity recognition models represents a seminal contribution to advancing the field's capabilities. The proposed design goes beyond a more comparative exercise by introducing novel methodologies and optimisations that enhance the practicality of indoor human activity recognition systems. Furthermore, the key outcome will be the ability to make these systems lighter and more computationally efficient. Moreover, the proposed design refines the theoretical underpinnings of indoor human activity recognition. It translates these advancements into tangible benefits, making the technology more accessible, adaptable, and relevant for real-time, lightweight, and practical applications. A comparison Table 2 summarises the main points from the search results related to multi-user WiFi-based human activity recognition.

Table 2 presents a comparative analysis of the methods, techniques, number of participants, limitations, accuracy, bandwidth, number of transmitters and receivers, and other factors used in various studies related to WiFi-based multi-human activity recognition. The table displays all of the proposed techniques for detecting human activity using WiFi signals. Furthermore, evaluating LOS and NLOS conditions is pivotal to understanding multi-user sensing within indoor environments. LOS scenarios, where a direct and unobstructed path exists between the transmitter and receiver, contrast sharply with NLOS situations, where obstacles such as walls or partitions disrupt the direct line of signal propagation. The proposed model's sophisticated signal processing techniques augment its capacity to discern and interpret signals in challenging NLOS conditions. The model improves activity recognition accuracy by leveraging temporal dependencies and contextual information, even in an obstructed direct line of sight. The design enhancement allows the model to surpass the limitations of prior approaches, providing a more resilient and accurate multi-user sensing capability, especially in NLOS scenarios.

G. MULTI LOCATIONS EVALUATION AND MATCHING CAPABILITY

The environment has direct effect in WiFi sensing, especially in scenarios involving multiple users using CSI. Complex environments with varying layouts, furniture, and materials can introduce multipath effects, where the wireless signals bounce off surfaces, creating multiple overlapping signals. This complicates the extraction of clean activity patterns from the CSI data. Additionally, the presence of multiple users adds further complexity as their movements interfere with each other, leading to more intricate signal variations. Additionally, noise, both from external sources such as electronic devices and internal system noise, which further degrades the CSI quality.

The pursuit of location-independent WiFi sensing represents a contribution step in overcoming the challenges posed by environmental effects on signal propagation in various settings. The proposed model contributes by introducing a way to match locations that reduce the impact of environmental factors. This makes the WiFi sensing system more resistant to change in its surroundings and eliminates environmental effects by making it easier to build WiFi sensing systems that

TABLE 2. Comparative analysis of WiFi-Based HAR works, highlighting methods, techniques, number of people, bandwidth, number of antennas, accuracy, and limitations.

Ref	Method	Technique	No. of People	BW	No. APs	Accuracy	Limitation
[22]	Multi-environment framework	WiFi CSI signals with an attention mechanism	10	20MHz	3	Achieved high accuracy in recognizing various	more users involved, increased interference, & decreased recognition performance
[20]	Respiration sensing using WiFi signals	Commodity WiFi	Multiple people	5.2GHz	9	Achieved high accuracy in multi-person respiration sensing	Limited to respiration sensing
[30]	Continuous action recognition	WiFi signals	eighteen	20MHz	3	Achieved high accuracy in multi-user continuous action recognition	Unable to recognize the number of people and the lack of current wall-penetrating capability
[11]	Presence and activity detection	5G	10	0.4GHz	2	Achieved high accuracy in multi-user presence and activity detection	Relies on 5G infrastructure, with limited spatial resolution & complexity in processing large connections
[34]	Tracking and Localization	mmWave	2	60-64GHz	2	Dopple effects provide high localization ability	Requires complex hardware and configuration and is unable to classify activities
[35]	Multi-users Gesture Classifier	Beamforming	2	600Mhz	2	Uses of spatial Beamforming to recognize hand gestures of multiple users	Beamforming requires a complex system and suffers from overlapping within close indoor coverage areas
This work	Presence and activity detection	WiFi signals	3	5.2GHz	2	Achieved high accuracy in multi-user continuous action recognition	Requires precise settings to match the original configuration

work seamlessly in a variety of settings with little need for location-specific calibration.

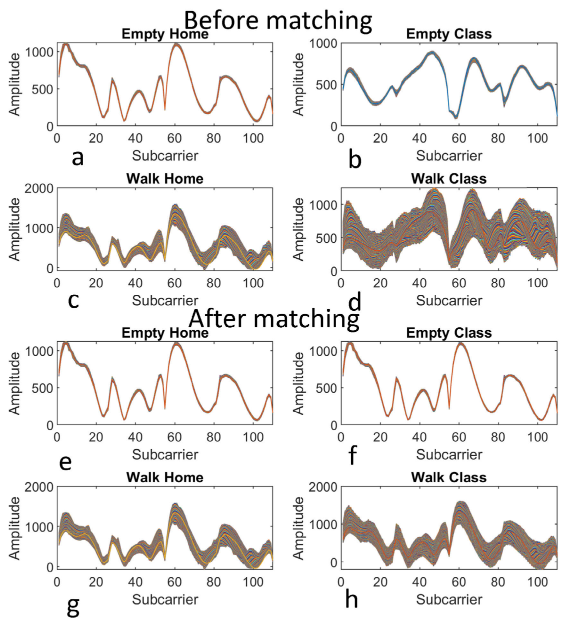

FIGURE 19. Variations across locations before matching (a, b, c, d) and after matching processes (e, f, g, h).

The activities in Fig. 18(a, b, c, d) highlight the contrasting spatial attributes between the two captured places before matching them. On the other hand, Fig. 19(e, f, g, h) demonstrates how the matching process adjusts the variances between locations for alignment. The evaluation of the designed model across multiple locations provides a thorough assessment of its performance in a variety of settings as shown in Fig. 16. The evaluation considers changes in signal propagation introduced by distinct ecological factors, such as architectural differences, material compositions, and spatial configurations. The technique reveals the model's adaptability to these variations, demonstrating its ability to recognize and classify human activities across locations independently and consistently.

H. LIMITATIONS AND FUTURE WORKS

The intricate temporal and spatial features used in this work of utilising LSTM and signal processing techniques contribute to a nuanced understanding of the signal dynamics. Furthermore, the attention mechanisms used in this model enables to focus on salient features, thus mitigating the impact of overlapping signals and increasing its discrimination capabilities. Additionally, incorporating such techniques underscores the system's adaptability to real-world complexities, setting the stage for further indoor human activity recognition improvement. Additionally, the potential for mis-classification, especially in scenarios with overlapping activities, necessitates ongoing refinement to reduce false positives and negatives. Therefore, the system must be improved for large-scale deployments, and its robustness in handling dynamic environmental conditions is another dimension.

The envisaged future works emanate from the identified limitations and aim to propel the system's capabilities to different environments. They need to address the complexity of human activities and investigate adaptive algorithms and learning mechanisms to adjust to changing conditions. Future research may also integrate real-time feedback mechanisms to enhance the system's adaptability. Further refinement of the model's interpretability enables a deeper understanding of its decision-making processes. Additionally, scalability studies and investigations into the system's performance in real-world scenarios with diverse activities will contribute to its applicability across broader contexts, such as through wall applications. Collaborative efforts with experts in architecture, signal processing, and human behaviour sciences may yield interdisciplinary insights, fostering the evolution of more sophisticated and practical indoor human activity recognition systems. One additional gap in this work is the

The MIMO systems used in this work employ clustored RPi at both the receiver to improve sensing performance, propagate signals using Omnidirectional antenna. While this omnidirectional propagation can enhance communication coverage and reliability, it poses a limitation for sensing applications, particularly in WiFi-based human activity recognition or localization, precise and controlled signal propagation. The Omnidirectional nature lead to multipath interference, where signals reflect off various surfaces, creating complex and unpredictable propagation patterns. This can reduce the accuracy and reliability of sensing measurements and can not distunglish between each individual's activity. In contrast, directional antennas focus the signal in a specific direction, reducing multipath effects and enhancing the precision of the received signals. By concentrating the signal in a targeted area, directional antennas, array antennas, and ultra band systems which can achieve better spatial resolution, making them more suitable for sensing applications that require high accuracy, such as detecting the exact position or movement of an object.

V. CONCLUSION

In conclusion, the developed multi-user sensing method advances location-independent multi-user indoor human activity recognition using WiFi signals. The proposed approach has enabled multi-user sensing, showcasing its robustness in discerning and classifying diverse human activities within complex, real-world environments. The strategic fusion of ICA and CWT contributes to the system's ability to capture intricate temporal and spatial features inherent in human movements in complex environments and for multi users. At the same time, the seq2seq LSTM algorithm enhances the model's contextual understanding, in particular in scenarios involving overlapping activities. This model's distinguishing strength lies in its ability to achieve remarkable accuracy in multi-user sensing while operating with a simplified hardware infrastructure specially with usage of attention mechanism. The minimalist hardware requirements make it readily deployable, offering a practical solution for various settings without necessitating extensive resources. The model achieves high accuracy in complex scenarios attests to its effectiveness in addressing the demands of dynamic environments with multiple users engaged in diverse activities. Furthermore, reducing layering complexity without compromising accuracy further highlights the system's efficiency and alterability. The groundbreaking culmination of this work holds immense promise for transformative advancements in real-world, multi-user sensing applications location independently.

REFERENCES

[1] J. Yang, H. Zou, H. Jiang, and L. Xie, "Device-free occupant activity sensing using WiFi-enabled IoT devices for smart homes," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3991–4002, Oct. 2018.

[2] S. Kim and J.-H. Yun, "Motion-aware interplay between WiGig and WiFi for wireless virtual reality," *Sensors*, vol. 20, no. 23, p. 6782, Nov. 2020.

[3] Y. Zhang, Y. Yin, Y. Wang, J. Ai, and D. Wu, "CSI-based location-independent human activity recognition with parallel convolutional networks," *Comput. Commun.*, vol. 197, pp. 87–95, Jan. 2023.

[4] X. Ding, T. Jiang, Y. Zhong, S. Wu, J. Yang, Z. Gu, T. He, Z. Wang, Y. Xu, and J. Yang, "Wi-Fi-based location-independent human activity," *Comput. Commun.*, vol. 71, pp. 1–17, Nov. 2022.

[5] P. F. Moshiri, R. Shahbazian, M. Nabati, and S. A. Ghorashi, "A CSI-based human activity recognition using deep learning," *Sensors*, vol. 21, no. 21, p. 7225, Oct. 2021.

[6] Y. Zhang, Y. Zheng, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, "Widar3.0: Zero-effort cross-domain gesture recognition with Wi-Fi," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 11, pp. 8671–8688, Nov. 2022.

[7] S. K. Yadav, S. Sai, A. Gundewar, H. Rathore, K. Tiwari, H. M. Pandey, and M. Mathur, "CSITime: Privacy-preserving human activity recognition using WiFi channel state information," *Neural Netw.*, vol. 146, pp. 11–21, Feb. 2022.

[8] Y. Zhang, Q. Liu, Y. Wang, and G. Yu, "CSI-based location-independent human activity recognition using feature fusion," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022.

[9] R. H. Venkatnarayan, G. Page, and M. Shahzad, "Multi-user gesture recognition using WiFi," in *Proc. 16th Annu. Int. Conf. Mobile Syst., Appl., Services*, Munich, Germany, Jun. 2018, pp. 401–413.

[10] R. Zhang, Y. Zhou, and B. Qu, "High resolution CSI feedback with beam space MIMO," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Bologna, Italy, Oct. 2017, pp. 1–5.

[11] A. M. Ashleibta, A. Taha, M. A. Khan, W. Taylor, A. Tahir, A. Zoha, Q. H. Abbasi, and M. A. Imran, "5G-enabled contactless multi-user presence and activity detection for independent assisted living," *Sci. Rep.*, vol. 11, no. 1, pp. 1–15, Sep. 2021.

[12] T. Gu, Z. Fang, Z. Yang, P. Hu, and P. Mohapatra, "MmSense: Multi-person detection and identification via mmWave sensing," in *Proc. 3rd ACM Workshop Millim.-Wave Netw. Sens. Syst.*, Los Cabos, Mexico, Oct. 2019, pp. 45–50.

[13] Y. He, Y. Chen, Y. Hu, and B. Zeng, "WiFi vision: Sensing, recognition, and detection with commodity MIMO-OFDM WiFi," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 8296–8317, Sep. 2020.

[14] S. Tan, L. Zhang, Z. Wang, and J. Yang, "MultiTrack: Multi-user tracking and activity recognition using commodity WiFi," in *Proc. CHI Conf. Human Factors Comput. Syst.*, May 2019, pp. 3–5.

[15] R. H. Venkatnarayan, S. Mahmood, and M. Shahzad, "WiFi based multi-user gesture recognition," *IEEE Trans. Mobile Comput.*, vol. 20, no. 3, pp. 1242–1256, Mar. 2021.

[16] L. S. Ambati and O. El-Gayar, "Human activity recognition: A comparison of machine learning approaches," *J. Midwest Assoc. Inf. Syst.*, vol. 2021, no. 1, pp. 1–13, 2021.

[17] S. Majumder, E. Aghayi, M. Noferesti, H. Memarzadeh-Tehran, T. Mondal, Z. Pang, and M. Deen, "Smart homes for elderly healthcare—Recent advances and research challenges," *Sensors*, vol. 17, no. 11, p. 2496, Oct. 2017.

[18] X. Zhu, W. Qu, X. Zhou, L. Zhao, Z. Ning, and T. Qiu, "Intelligent fingerprint-based localization scheme using CSI images for Internet of Things," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 4, pp. 2378–2391, Jul. 2022.

[19] Q. Li, R. Gravina, Y. Li, S. H. Alsamhi, F. Sun, and G. Fortino, "Multi-user activity recognition: Challenges and opportunities," *Inf. Fusion*, vol. 63, pp. 121–135, Nov. 2020.

[20] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, "MultiSense: Enabling multi-person respiration sensing with commodity WiFi," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 4, no. 3, pp. 1–29, Sep. 2020.

[21] D. Wang, J. Yang, W. Cui, L. Xie, and S. Sun, "AirFi: Empowering WiFi-based passive human gesture recognition to unseen environment via domain generalization," *IEEE Trans. Mobile Comput.*, vol. 23, no. 2, pp. 1156–1168, Feb. 2022.
[22] X. Duan, Y. Zhang, Y. Li, and S. Li, "Wisdom of the crowd-based activity recognition in smart homes," *J. Ambient Intell. Humanized Comput.*, vol. 14, pp. 133–143, Jan. 2023.
[23] P. Wang, B. Guo, T. Xin, Z. Wang, and Z. Yu, "TinySense: Multi-user respiration detection using Wi-Fi CSI signals," in *Proc. IEEE 19th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, Dalian, China, Oct. 2017, pp. 1–6.
[24] S. Depatla and Y. Mostofi, "Crowd counting through walls using WiFi," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Athens, Greece, Mar. 2018, pp. 1–10.
[25] J. Liu, L. Wang, J. Fang, L. Guo, B. Lu, and L. Shu, "Multi-target intense human motion analysis and detection using channel state information," *Sensors*, vol. 18, no. 10, p. 3379, Oct. 2018.
[26] F. Abuhoureyah, W. Y. Chiew, and M. S. Zitouni, "WiFi based human activity recognition using multi-head adaptive attention mechanism," *J. Intell. Fuzzy Syst.*, pp. 1–16, Apr. 2024.
[27] X. Dang, X. Si, Z. Hao, and Y. Huang, "A novel passive indoor localization method by fusion CSI amplitude and phase information," *Sensors*, vol. 19, no. 4, p. 875, Feb. 2019.
[28] X. Ding, T. Jiang, Y. Zhong, Y. Huang, and Z. Li, "Wi-Fi-based location-independent human activity recognition via meta learning," *Sensors*, vol. 21, no. 8, p. 2654, Apr. 2021.
[29] F. S. Abuhoureyah, Y. C. Wong, and A. S. B. M. Isira, "WiFi-based human activity recognition through wall using deep learning," *Eng. Appl. Artif. Intell.*, vol. 127, Jan. 2024, Art. no. 107171.
[30] J. He and W. Yang, "IMar: Multi-user continuous action recognition with WiFi signals," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 6, no. 3, pp. 1–27, Sep. 2022.
[31] J. Yu, P. Wang, T. Koike-Akino, Y. Wang, P. V. Orlik, and R. M. Buehrer, "Multi-band Wi-Fi sensing with matched feature granularity," *IEEE Internet Things J.*, vol. 9, no. 23, pp. 23810–23825, Dec. 2022.
[32] B. Korany, H. Cai, and Y. Mostofi, "Multiple people identification through walls using off-the-shelf WiFi," *IEEE Internet Things J.*, vol. 8, no. 8, pp. 6963–6974, Apr. 2021.
[33] L. Guan, Z. Zhang, X. Yang, N. Zhao, D. Fan, M. A. Imran, and Q. H. Abbasi, "Multi-person breathing detection with switching antenna array based on WiFi signal," *IEEE J. Transl. Eng. Health Med.*, vol. 11, pp. 23–31, 2023.
[34] W. Li, Y. Wu, R. Chen, H. Zhou, and Y. Yu, "Indoor multihuman device-free tracking system using multiradar cooperative sensing," *IEEE Sensors J.*, vol. 23, no. 22, pp. 27862–27871, Nov. 2023.
[35] R. Peng, Y. Tian, and S. Han, "Multiuser wireless hand gesture recognition by spatial beamforming," *IEEE Trans. Veh. Technol.*, vol. 72, no. 1, pp. 805–820, Jan. 2023.
[36] L. Zhang, Q. Gao, X. Ma, J. Wang, T. Yang, and H. Wang, "DeFi: Robust training-free device-free wireless localization with WiFi," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8822–8831, Sep. 2018.
[37] J. Wang and J. Park, "An enhanced indoor positioning algorithm based on fingerprint using fine-grained CSI and RSSI measurements of IEEE 802.11n WLAN," *Sensors*, vol. 21, no. 8, p. 2769, Apr. 2021.
[38] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free your CSI: A channel state information extraction platform for modern Wi-Fi chipsets," in *Proc. 13th Int. Workshop Wireless Netw. Testbeds, Experim. Eval. Characterization*, Oct. 2019, pp. 21–28.
[39] D. Konings, F. Alam, F. Noble, and E. M.-K. Lai, "Device-free localization systems utilizing wireless RSSI: A comparative practical investigation," *IEEE Sensors J.*, vol. 19, no. 7, pp. 2747–2757, Apr. 2019.
[40] M. Kciuk, "OpenWRT operating system based controllers for mobile robot and building automation system students projects realization," in *Proc. 15th Int. Workshop Res. Educ. Mechatronics (REM)*, Sep. 2014, pp. 1–4.
[41] M. Atif, S. Muralidharan, H. Ko, and B. Yoo, "Wi-ESP—A tool for CSI-based device-free Wi-Fi sensing (DFWS)," *J. Comput. Design Eng.*, vol. 7, no. 5, pp. 644–656, Oct. 2020.
[42] Z. U. A. Akhtar and H. Wang, "WiFi-based driver's activity monitoring with efficient computation of radio-image features," *Sensors*, vol. 20, no. 5, p. 1381, Mar. 2020.
[43] Y. Zeng, D. Wu, J. Xiong, and D. Zhang, "Boosting WiFi sensing performance via CSI ratio," *IEEE Pervas. Comput.*, vol. 20, no. 1, pp. 62–70, Jan. 2021.

FAHD ABUHOUREYAH received the B.Sc. and Engineering degrees in electrical and electronics engineering from the University of Bradford, U.K., and the master's and Ph.D. degrees in electronic engineering (electronic systems) from Universiti Teknikal Malaysia Melaka. His current research interests include computer systems and their intersection with innovation studies, science and technology, and electronic devices.

KOK SWEE SIM (Senior Member, IEEE) is currently a Professor with Multimedia University, Malaysia. He is also working closely with various local and overseas institutions and hospitals. He has filed more than 18 patents and 70 copyrights. He is also a fellow of The Institution of Engineers, Malaysia (IEM), and Institution of Engineering & Technology (IET). He has received many international and local awards. He was a recipient of Japan Society for the Promotion of Science (JSPS) Fellowship, Japan, in 2018, the Top Research Scientists Malaysia (TRSM) from the Academic Science Malaysia, in 2014, Korean Innovation and Special Awards, in 2013, 2014, and 2015, and the TM Kristal Award and International Championships of World Summit on the Information Society (WSIS) Prizes, in 2017, 2018, 2019, 2020, and 2021.

YAN CHIEW WONG received the Doctorate degree from The University of Edinburgh, U.K., in 2014. She is currently an Associate Professor with Universiti Teknikal Malaysia Melaka (UTeM). Previously, she was with Infineon Technologies, Intel, and Sofant Technologies. She has more than eight years of industry experience in semiconductor design. She has intensively researched integrated circuit (IC) design, artificial intelligence (AI), and energy harvesting. She receives national and international awards, such as ITEX, INNOVATE, IEM, and EDS Societies. She is equipped with both design and implementation skills.