Developments in the Built Environment 23 (2025) 100745

Contents lists available at ScienceDirect

Developments in the Built Environment

journal homepage: www.sciencedirect.com/journal/developments-in-the-built-environment

A mobile receiver WiFi-CSI approach for fall detection of construction workers

Yinong Hu [a], Heng Li [a], Mingzhou Cheng [a,*], Mingyu Zhang [a], Shuai Han [a], Waleed Umer [b]

[a] Department of Building and Real Estate, Faculty of Construction and Environment, Hong Kong Polytechnic University, Hong Kong, China
[b] Department of Architecture and Built Environment, Northumbria University, NE1 8ST, Newcastle Upon Tyne, United Kingdom

ARTICLE INFO

Keywords:
Fall detection
Channel state information (CSI)
Smartphone
Mobile receiver
Doppler frequency
Construction safety

ABSTRACT

This study introduces a novel fall detection method for construction workers that uses WiFi Channel State Information (CSI) with mobile smartphone receivers, which addresses the high incidence of fall-related injuries at construction sites. The innovative approach utilizes Doppler frequency shift features captured through mobile receivers, which adapt to dynamic construction environments where workers continuously move, overcoming limitations of conventional static configurations. Our framework extracts characteristic CSI patterns from WiFi signals and employs an improved deep learning model to classify falls and common construction activities. Experimental validation demonstrates robust performance with accuracy exceeding 93 % across various distances and orientations. The mobile receiver design significantly enhances spatial adaptability while providing a non-invasive, privacy-preserving, and cost-effective solution that can be readily deployed using existing WiFi infrastructure and workers' smartphones for construction site safety monitoring.

1. Introduction

The construction industry is globally recognized as one of the most hazardous sectors. According to the International Labour Organization (ILO) (Occupational Safety and Health Management, 2021), an estimated 60,000 fatal accidents occur on construction sites worldwide each year, accounting for approximately one in six of all fatal workplace incidents. This alarming statistic is starkly illustrated in the United States, where the Occupational Safety and Health Administration (OSHA) reports that "Fall Protection" has been the most frequently cited violation for 13 consecutive years, with 7271 citations in 2023 alone (Schirn, 2024). According to a report from the Hong Kong Labour Department (Bulletin 2022_issue23_sc 2022), the construction industry consistently records the highest number of fatalities and accident rates among all sectors. Within these incidents, slipping, tripping, or falls on the same level, in particular, are the leading cause of worker accidents, accounting for 22 % of the total accidents (OSH_Statistics_2022_tc 2022). Given the aging workforce in many regions, a significant number of middle-aged and elderly workers are engaged at construction activities (Poon, 2009), falls pose a greater risk to aging workers, leading to injuries such as sprains, fractures, and musculoskeletal issues (Courtney et al., 2001; Lipscomb et al., 2006) that become increasingly life-threatening for this demographic. Therefore, the development of a system that can promptly monitor falls and provide feedback to management for timely treatment is crucial for ensuring worker safety.

In previous studies, different technologies and devices have been employed to monitor fall events on construction sites. These can be broadly categorized into three types: wearable-based systems, non-wearable-based systems, and fusion or hybrid-based systems (Chaccour et al., 2017). The wearable-based approach is a common choice. These sensors can be worn on various body parts of the workers, such as the wrist, waist, or the soles of the feet to detect and analyze body movements and postures (Antwi-Afari et al., 2018; Ferreira De Sousa et al., 2022; Choo et al., 2023a). This strategy, however, requires workers to wear additional devices, which can be inconvenient during work processes and may require frequent charging or battery replacement. Non-wearable methods involve deploying sensors in the ambient environment, such as pressure, passive infrared, vibration, acoustic, and infrared sensors (Guan et al., 2017; Alwan et al., 2006). Unfortunately, the expense of developing and maintaining these sensor systems across the construction site is a significant issue. Vision-based approaches also constitute a commonly used non-wearable detection method, enabling fall event detection through analysis of body posture and motion patterns in video images (Duan et al., 2023a; Fang et al., 2020; Ding et al.,

* Corresponding author.
E-mail address: 23123842r@connect.polyu.hk (M. Cheng).

https://doi.org/10.1016/j.dibe.2025.100745
Received 12 May 2025; Received in revised form 28 July 2025; Accepted 26 August 2025
Available online 28 August 2025

2018). However, vision-based monitoring may raise privacy concerns and require proper installation and maintenance. Additionally, at construction site scenes, objects often obstruct the view, leading to challenges in accurate image recognition (Fang et al., 2019a). Fusion-based methods involve integrating wearable and non-wearable approaches, utilizing multi-channel data sources with fusion algorithms to enhance the reliability and specificity of the detection system (M et al., 2023a; Nyan et al., 2006). This fusion approach also faces substantial computational demands, complex processing strategies, and elevated costs. In addition, some studies have explored fall detection using the built-in inertial measurement unit (IMU) sensors in smartphones, utilizing acceleration and angular velocity measurements (Dzeng et al., 2014; Fang and Dzeng, 2014). This detection performance, however, is highly dependent on the phone's orientation, requiring it to maintain a correct position for accurate detection (Santoyo et al., 2018).

CSI (Channel State Information) is a measurement technique used in wireless communication systems to gather information about the channel conditions between a transmitter and receiver. It provides non-invasive, wide coverage fall detection on construction sites, surpassing the limitations of wearable and non-wearable sensors, reducing equipment costs. Meanwhile, it offers flexibility, convenience, and rich data analysis capabilities. The potential of CSI signals for fall detection and other applications involving human activity recognition has been extensively investigated by researchers in recent years. Previous studies on utilizing CSI signals for human activity recognition primarily used fixed transmitters consisting of one or two laptops equipped with Intel 5300 WiFi cards and a router (Yan et al., 2022; Shalaby et al., 2022; Lee et al., 2020a). Wang et al. (2017a) reported that as individuals walked away from the straight-line distance between the receiver and transmitter, the model's accuracy declined. Accurate detection in fixed setups demands the deployment of multiple devices to cover the entire region. Furthermore, these studies focused on interior scenarios and identifying falls in the elderly, which poses challenges when applied to dynamic construction sites that are open, complicated, and susceptible to change (Al- et al., 2016). Construction workers' larger activity areas and more vigorous movements render previous models unsuitable. Deploying multiple fixed receivers for comprehensive coverage in diverse construction sites is challenging and may interfere with workers' activities, increasing equipment costs.

Based on the limitations and issues identified in previous research, this study aims to provide a convenient, flexible, and reliable method for fall detection on construction sites while leveraging existing WiFi infrastructure and smartphones. Unlike earlier studies, we propose a new method of combining WiFi and smartphones without the need for additional devices or specific WiFi cards. Through our custom-developed application, CSI data is wirelessly extracted from WiFi signals and directly stored on smartphones, offering flexibility and mobility. This approach offers significant advantages over solutions that rely only on the smartphone's built-in IMU sensor, as it does not rely on the phone's orientation or location. The main contributions of this research include developing a mobile receiver approach specifically adapted to the dynamic working environment of construction sites, eliminating the need for fixed device positioning. The system is low-cost and easy to deploy, reducing implementation costs by utilizing existing WiFi infrastructure and smartphones. Regarding construction management, the solution provides project managers with a tool to monitor worker safety and can be used to identify high-risk areas and activities, facilitating the improvement of site safety.

The remainder of this paper is organized as follows. Section 2 provides a comprehensive review of the relevant literature, examining existing fall detection methodologies and technologies. We then delineate the proposed research methodology in Section 3, covering the key stages from CSI extraction to the development of our detection model. In Section 4, a thorough evaluation of the system's performance is presented. Subsequently, Section 5 discusses the broader implications of our findings, acknowledges the limitations of this study, and proposes directions for future research. Finally, Section 6 concludes the paper with a summary of our contributions.

2. Literature review

2.1. Conventional sensing methods

Wearable devices represent a predominant method for fall detection in the construction industry. Their primary function is to identify potential fall risks by measuring workers' Loss of Balance (LoB). These devices can be divided into accelerometers, gyroscopes, or IMUs (Ferreira De Sousa et al., 2022; Kim et al., 2020; Yang et al., 2015) worn near the body's center of gravity (such as the waist or head) and devices installed on shoe soles or feet to monitor gait stability (Antwi-Afari et al., 2020; Lin et al., 2022). For instance, Kim et al. (2020) developed an IMU system that achieves this level of accuracy by monitoring sudden posture changes in workers. Similarly, Maxwell et al. (Antwi-Afari et al., 2020) designed a smart insole pressure system that can predict unstable states by analyzing abnormalities in workers' gait. However, wearable devices face numerous challenges in practical applications. The significant differences in individual workers' movement patterns make data interpretation difficult and prone to misjudgments (Hong et al., 2023). Improper sensor positioning directly affects detection accuracy, leading to unreliable results. Moreover, construction workers typically exhibit reluctance to wear devices for extended periods, with research surveys indicating that approximately 54 % of field workers are unwilling to voluntarily use biometric wearable sensing devices, while a higher 59 % refuse to use location-tracking wearable devices (Fugate and Alzraiee, 2023). Additionally, the limitations in battery life and the exposure of devices to dust, moisture, and impacts present at construction environments have the potential to compromise their reliability.

Non-wearable fall detection techniques were less intrusive than wearable technology. The effectiveness of using non-wearable technologies, including cameras or ambient sensors, like vibration or sound detectors, to record the surrounding environment and track workers' movements has been shown in some studies (Lee et al., 2020b). Due to the complicated environments, the investigation into visual-based detection techniques has been more thorough. For example, Duan et al. (2023b)utilized body posture stability as a means to identify worker activities, while Liu et al. (2023)developed a computer vision-based non-contact fall detection method to detect fall accidents among construction workers. Although non-wearable fall detection methods offered advantages in terms of invasiveness, they also presented certain challenges and limitations. Cameras have monitoring blind spots, making it difficult to comprehensively cover complex worksites. Changes in lighting, obstructions, and dust at construction sites significantly affect system reliability (Fang et al., 2019b). Continuous video monitoring also raises worker privacy concerns (Seo et al., 2015). Additionally, deploying high-definition camera networks and processing real-time video requires substantial computational resources and investment.

Multi-modal fusion methods attempt to overcome the limitations of single technologies by combining multiple sensing techniques. Amsaprabhaa et al. (M et al., 2023b) improved detection accuracy by integrating visual and skeletal kinematic features, while Choo et al. (2023b) developed a high-altitude work safety monitoring system by integrating IMUs and barometric pressure sensors. Multi-modal methods demonstrate significant advantages in accuracy and environmental adaptability but also increase system complexity and cost. Data synchronization issues, deployment difficulties, and high maintenance costs make these methods challenging to implement in practical applications.

2.2. WiFi-based activity recognition and fall detection

Compared to the conventional fall detection methods discussed in

Section 2.1, WiFi-based solutions provide significant advantages. These systems can utilize the existing network infrastructure, thereby substantially decreasing implementation expenses. Additionally, WiFi technology does not collect personally identifiable information, protecting worker privacy and security. Crucially, environmental lighting conditions do not affect WiFi system performance, allowing reliable operation in adverse weather and low visibility conditions. These characteristics collectively make WiFi-based solutions particularly suitable for dynamic and complex environments like construction sites.

CSI is a refined measurement and description by the receiver of the characteristics such as attenuation, phase shift, and multipath effects that occur in space and time to the wireless signals sent by the transmitter during WiFi communication. Researchers have validated the feasibility of CSI for activity recognition across multiple domains. In healthcare, CSI has been utilized to monitor elderly individuals' activities and fall status by analyzing fluctuations in wireless signals, thereby identifying actions such as sitting, lying down, and walking, which offer essential information for patient health assessment (Tan et al., 2018). In human-computer interaction, CSI has been employed to recognize gestures and body movements (P and Udgata, 2024), with researchers developing techniques capable of identifying specific gestures by analyzing wireless signals resulting from hand movements (Tang et al., 2021; Thariq et al., 2020). Furthermore, research has demonstrated that CSI can penetrate walls and obstacles for human activity recognition (Abuhoureyah et al., 2024). These research investigations collectively validate the feasibility and extensive application potential of CSI in activity recognition.

In the construction field, Guo et al. (2023) made a pioneering contribution by applying WiFi channel state information to construction worker fall detection. However, their experimental design followed the conventional treatment in other fields of using two fixed computers as transmitting and receiving devices. This arrangement exhibits multiple challenges in changing construction site conditions. Initially, the deployment of numerous fixed devices over a construction site presents logistical complexities and time-consuming tasks, necessitating careful preparation and installation to guarantee sufficient coverage and overlap among devices. This approach necessitates that the worker be situated between the two devices, which may restrict movement and impact work efficiency. Additionally, the detection accuracy of CSI with fixed receivers is significantly affected by the distance from the line-of-sight (LoS) path. In wireless communications, the LoS path refers to the direct propagation path from transmitter to receiver, and when workers move away from this direct path, the model's detection sensitivity typically decreases. Wang (Wang et al., 2017a) established that detection rates decline markedly with increasing distance from the LoS path. At construction site environments, maintaining a stable line-of-sight propagation path is extremely difficult due to numerous obstacles and constantly changing equipment layouts, posing serious challenges for fixed receiver configurations in practical applications.

To address these limitations, our research takes an innovative approach by introducing mobile receiving devices (such as smartphones) for fall detection at construction sites. This approach offers significant advancement over conventional fixed-device configurations. Workers carrying the receiving device ensure that the monitored subject always remains close to the LoS path between transmitter and receiver, significantly enhancing detection sensitivity. This adaptive approach diminishes dependence on fixed devices, simplifies implementation, and improves the overall flexibility of fall detection methods. Utilizing the widespread adoption of smartphones, our solution offers a cost-efficient and easily accessible method for fall detection on construction sites, while effectively overcoming the problem of reduced detection sensitivity caused by workers staying away from LoS paths in conventional fixed equipment configurations.

However, employing a mobile receiver also introduces a range of technical challenges, such as Doppler effects, difficulty distinguishing between signal variations caused by receiver movement and those caused by target actions, and increased signal processing complexity in multipath environments. These challenges, along with their corresponding solutions, will be discussed in detail in the methodology and discussion sections.

3. Methodology

3.1. Preliminaries

Wireless signals are influenced by objects and human activities in the environment during propagation. When a human body moves near a wireless transmission path, the signal experiences multipath effects including reflection, scattering, and diffraction, causing variations in the amplitude and phase of the received signal. The channel's changing amplitudes and phases are a reflection of the variations in the signal as it moves through it, which match the variations in the movement of the body's parts (Yousefi et al., 2017; Qian et al., 2017). This is why we decided to perform the analysis of the fall experiment by extracting the CSI.

In the current mainstream 802.11 WiFi systems, based on Orthogonal Frequency-Division Multiplexing (OFDM) technology, the data stream is divided into multiple substreams, which are transmitted in parallel through multiple subchannels of different frequencies (Patil et al., 2017). Each subchannel utilizes independent subcarriers, and CSI is the channel attribute used to describe each subchannel in this system, specifically referring to the values of each element in the channel matrix H. The channel matrix can be shown as follows, Eq. (1), assuming that the transmitter has M antennas, the receiver has N antennas, and that K subcarriers are utilized for communication:

$$\mathrm{CSI}_i = \begin{bmatrix} \mathrm{H}_{1,1} & \cdots & \mathrm{H}_{1,N} \\ \vdots & \vdots & \vdots \\ \mathrm{H}_{j,1} & \mathrm{H}_{j,k} & \mathrm{H}_{j,N} \\ \vdots & \vdots & \vdots \\ \mathrm{H}_{M,1} & \ldots & \mathrm{H}_{M,N} \end{bmatrix} \tag{1}$$

where *i* represents the *i-th* CSI packet transmission. Each element $\mathrm{H}_{j,k}$ in the channel matrix represents the channel response between the *j-th* transmit antenna and the *k-th* receive antenna. It can be represented as Eq. (2):

$$\mathrm{H}_{j,k} = \left[\mathrm{h}_{j,k,1}, \mathrm{h}_{j,k,2}, \ldots, \mathrm{h}_{j,k,K}\right] \tag{2}$$

Multipath signal propagation appears as time-delay spread in the time domain and causes frequency-selective fading in the frequency domain. Therefore, CSI can be described through Channel Frequency Response (CFR), which characterizes multipath signal propagation through amplitude-frequency and phase-frequency characteristics. CFR can be represented as Eq. (3):

$$\mathrm{h}_{j,k,K} = \left|\left|\mathrm{h}_{j,k,K}\right|\right| e^{j\angle \mathrm{h}_{j,k,K}} \tag{3}$$

where $\mathrm{h}_{j,k,K}$ represents the Channel State Information of the *k-th* subcarrier, $\left|\left|\mathrm{h}_{j,k,K}\right|\right|$ represents the amplitude of the *k-th* subcarrier, and $\angle \mathrm{h}_{j,k,K}$ represents the phase of the *k-th* subcarrier.

CFR has a direct and distinct correlation with human activities. Different activities produce specific motion patterns that affect wireless signal propagation paths in unique ways. These characteristic patterns in CFR data can be used to distinguish various activity states. More specifically, by analyzing CFR amplitude, phase change rates, and inter-subcarrier correlations, we can extract key features that differentiate normal activities from fall events.

The Intel 5300, a widely used CSI hardware, can access 30 subcarriers. However, the number of accessible subcarriers has surged to about 256 with the introduction of Nexmon devices that support 802.11ac and 80 MHz channels. This expanded access to subcarriers can

potentially improve the accuracy and reliability of activity recognition models that rely on CSI data. The acquired CSI data demonstrates that the CSI amplitude fluctuation in a 20 MHz bandwidth under 2.4G band is more noticeable with the movement of the human body, hence in this experiment, we perform the data acquisition in a 20 MHz bandwidth, which can obtain 64 subcarriers. The experimental setup is to select one antenna of the router as the transmitting side, while the smartphone at the receiving side has only one antenna. Therefore, the collected CSI data can be represented as Eq. (4) and Eq. (5):

$$H = \left[h_{1,1,1}, h_{1,1,2}, h_{1,1,3} \ldots h_{1,1,64}\right] \tag{4}$$

$$h_{1,1,1} = [H(t_1) \cdots H(t_n)] \tag{5}$$

H represents the CSI data acquired at this setting, where each H(t_n) represents the CFR on that subcarrier at time point t_n, recording the amplitude and phase information of the signal at that moment.

3.2. Experiment design

3.2.1. Experiment setup

Since the release of Linux 802.11n CSI tools (Halperin et al., 2011), the majority of commonly used receivers are paired with systems equipped with Intel 5300 wireless cards. A laptop that can support this wireless card is needed for the receiver. With the release of the Nexmon firmware (Schulz et al., 2016), it is now possible to extract CSI from mobile phones. Compared to conventional CSI collection methods based on Intel 5300 wireless cards, Nexmon allows users to modify and customize wireless network card firmware, enabling the capture of more comprehensive and detailed CSI data on smartphones, including channel state, multipath effects, interference, and signal strength.

In our experimental setup, we selected a TL-XDR3010 WiFi router as the transmitter and a Nexus 6p smartphone as the receiver. The Nexus 6p was chosen because it features a Broadcom WiFi chipset (bcm4358) compatible with Nexmon firmware. The router operated at 2.4 GHz frequency on channel 13, employing MIMO mode with OFDM transmission. Using the PSPing tool to generate network traffic, the router broadcasts approximately 100 data packets per second on average.

To facilitate efficient data collection, we developed an application using Android Studio to run on the Nexus 6P. This application controls and manages the WiFi channel state information acquisition process, storing data in the smartphone's memory. The application invokes Nexmon firmware to extract CFR matrices, with configuration parameters including channel 13 and 20 MHz bandwidth. The design ensures receiver mobility while enabling real-time data storage and convenient Nexmon tool integration, eliminating the need for data cable connections during collection. Through our custom program, each dataset can collect information from 64 subcarriers.

Fig. 1 illustrates the communication mechanism between components in the experimental system. The laptop generates traffic by sending ping packets to the WiFi router, which responds with pong packets. The smartphone then captures CSI at each pong packet transmission, recording physical environment interference information between the smartphone and router, which reflects the impact of human activities on wireless signals.

3.2.2. Data collection

Our data collection was conducted in an outdoor environment designed to simulate the complex conditions of a construction site. To introduce realistic dynamic interference, the site was chosen to be adjacent to a busy main road, ensuring a constant and unpredictable background of signal fluctuations from passing vehicles and pedestrians. Furthermore, during the experiments, non-participating personnel would occasionally walk through the periphery of the test area, adding another layer of human-induced dynamic noise. The environment also included static sources of interference typical of a worksite, such as numerous stacks of construction materials and continuously operating HVAC and plumbing systems, which created a rich multipath environment. This setup ensured that our model was trained and tested against a backdrop of realistic, uncontrolled environmental noise rather than in a static setting.

The data collection design intends to construct a robust and generalizable model by adding diversity to the dataset. We established three distinct horizontal distances of 2m, 6m, and 10m between the WiFi transmitter and the mobile receiver to evaluate performance under different spatial configurations. Participants were instructed to perform their activities while ensuring the smartphone, as the mobile receiver, maintained an approximate distance to the WiFi transmitter. To simulate realistic usage, the smartphone's placement was also varied, being positioned in the participant's hand or their trouser pocket, as illustrated in the experimental scenes in Fig. 2. Furthermore, we incorporated orientational diversity, conceptualized in Fig. 3, by having participants perform activities in distinct, representative orientations relative to the transmitter. This approach ensures the model learns the intrinsic features of an activity, a process aided by our mobile-receiver architecture, where the subject's body is always the primary scatterer near the receiver, rather than learning dependencies on specific geometric configurations.

The participants performed five activities to test the capability of the method to differentiate between falls and other vigorous movements. We selected squatting, brick lifting, and climbing as they are common activities within a construction environment. These movements represent a class of challenging, fall-like activities of daily living (ADLs). The literature widely confirms that such ADLs are a primary source of false alarms (Wang et al., 2017a, 2017b), as their large vertical motions produce CSI fluctuations similar to a genuine fall. The fall event was simulated as a rapid, large-scale body motion involving a sudden loss of balance from a standing or walking posture and a collapse onto a safety mat. Other fall-like activities included squatting, which involved a vertical displacement of approximately 0.8–1 m, and brick lifting, where participants bent at the knees and waist to pick up a standard brick

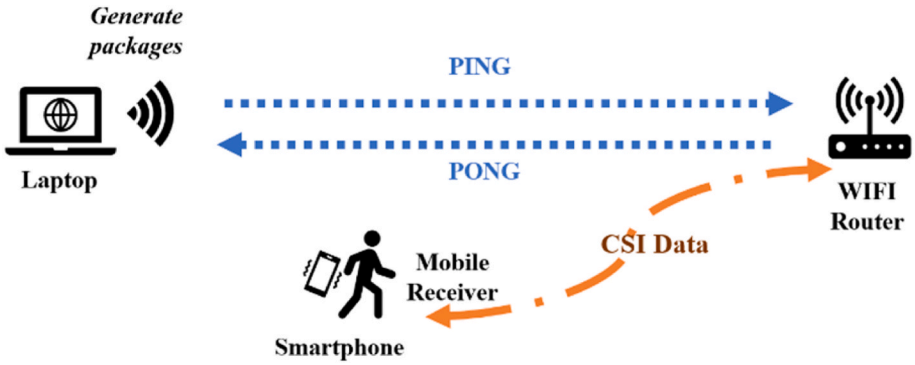

Fig. 1. Activity capture layout.

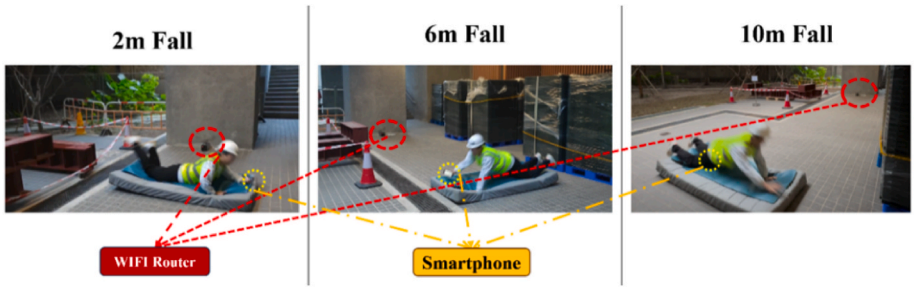

Fig. 2. Experimental scenes.

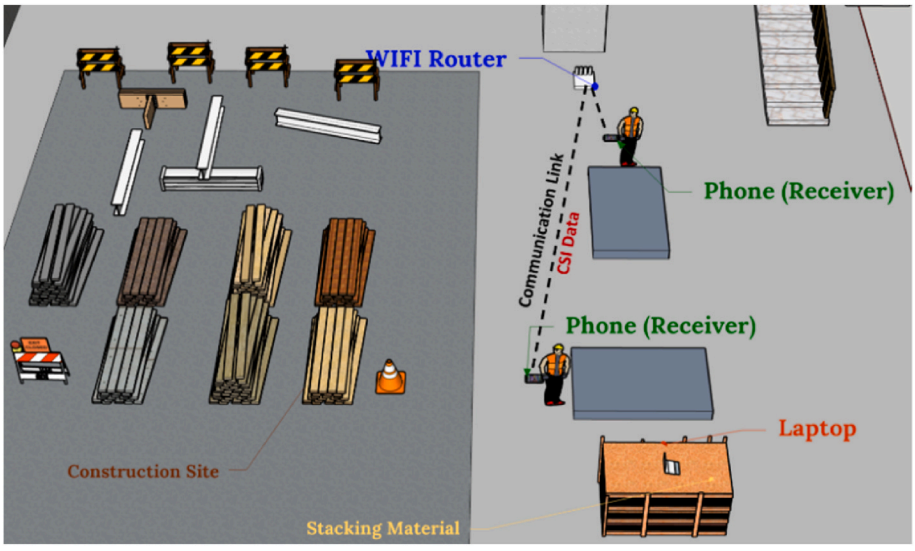

Fig. 3. 3D model of the experiment environment.

before returning to an upright stance. Ladder climbing and stair climbing were designed as continuous motions of ascending or descending a three-step ladder or a flight of stairs, respectively, all within the 5-s collection window. All these postures involve changes in the vertical position of the human body. The demographic information of the six university students recruited for this study is detailed in Table 1. While students' execution of specific tasks may differ from seasoned workers, the fundamental biomechanics of an involuntary fall are expected to be largely consistent.

Each participant completed a total of 120 sets of experiments, including 50 fall tests and 70 fall-like common construction activities. The fall tests were performed at three different distances to assess the impact of distance on detection effectiveness. All six participants completed a total of 720 sets of experimental data. Due to the influence of network topology on the speed of generating traffic in WiFi, the flow rate is not uniform. Each dataset lasted approximately 5 s, automatically stopping after collecting 500 packets, with each packet containing information from 64 subcarriers, creating a 500 × 64 data sample size. A researcher simultaneously recorded the experiments on video for precise data annotation.

Table 1
Demographic information of participants.

Participants	Age (Years)	Height (cm)	Weight (kg)	Gender
1	24	171	57	Female
2	30	167	55	Female
3	26	183	60	Male
4	23	182	86	Male
5	24	180	90	Male
6	22	185	95	Male

All experimental procedures were approved by the Human Subjects Ethics Sub-committee of The Hong Kong Polytechnic University (PolyU) under the reference number HSEARS20231016005. All participants provided informed written consent before the experiments and were informed that they could withdraw at any time.

3.3. Data analysis

This section provides a detailed description of the complete processing workflow from raw WiFi-CSI data to feature representations for fall detection. Fig. 4 illustrates the entire data processing framework, encompassing two key stages: data preprocessing and Doppler spectral feature extraction.

3.3.1. Preprocessing

The primary goal of data preprocessing is to transform raw WiFi-CSI data into high-quality channel information that accurately reflects the impact of human activities on wireless signals. Our preprocessing workflow was specifically optimized for CSI data collected from mobile receivers.

During data collection, the transmitter sent UDP packets at a predefined rate, while the receiver captured these packets using modified Nexmon firmware. Each CSI sample contained a complex channel frequency response across OFDM subcarriers. We extracted UDP packets from.pcap files and employed appropriate unpacking methods for the bcm4358 WiFi chipset to convert the data into complex form with real and imaginary parts. This resulted in a CFR matrix with dimensions (500,64), containing amplitude and phase information for 64

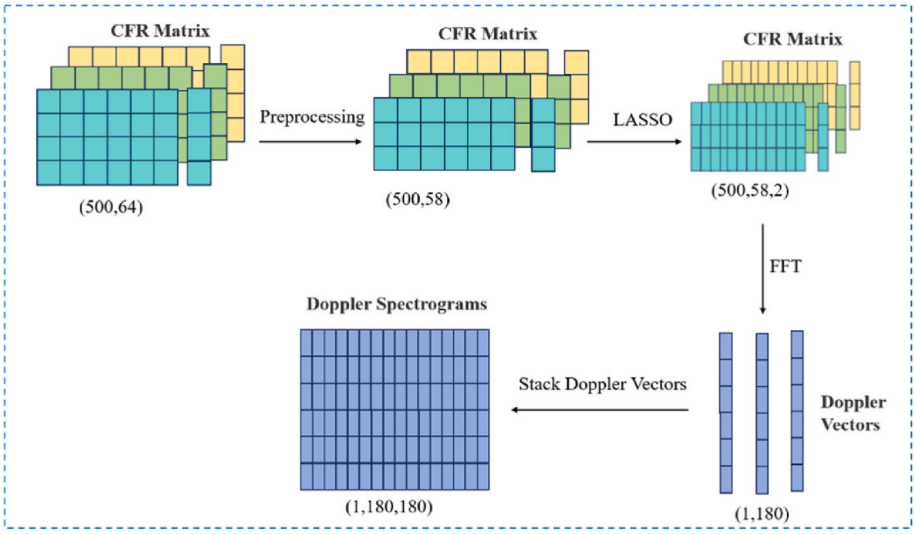

Fig. 4. CSI data processing framework.

subcarriers across 500 packets with millisecond-level timestamps, stored in.mat format.

CSI data in mobile receiver scenarios faces several unique challenges. When the receiving device moves, signal propagation paths continuously change, leading to more complex multipath effects. Receiver movement also introduces significant Doppler effects, causing signal frequency shifts (Sarkar et al., 2003). Additionally, channel characteristics during movement exhibit time-varying properties (Sayeed and Aazhang, 1999), making sampling unstable. Due to hardware limitations, smartphones typically generate CSI data with severe phase jumps and noise interference (Wang and Ho, 2024; Liu et al., 2021). These factors impact CSI data quality significantly, particularly phase information, reducing the accuracy of human activity recognition.

We implemented a comprehensive signal processing strategy based on the SHARP framework (Meneghello et al., 2023). The raw CSI data is first processed by FFT spectral shift (fftshift) to place the zero frequency components at the center of the spectrum for subsequent analysis. A crucial subsequent step in the processing workflow involved the elimination of invalid data, primarily targeting zero-value rows and specific subcarriers (such as DC subcarriers and edge subcarriers), which typically carry noise rather than effective signals. Each of the retained subcarriers is amplitude-normalized to reduce the effect of different received strengths on the results. After these steps, the raw CSI data is transformed into normalized and cleaned frequency domain data, which establishes the foundation for subsequent processing.

To convert the original frequency domain CSI data into a high-precision time domain representation, we employed time domain high-resolution reconstruction techniques. This step uses the LASSO regression algorithm (Ranstam and Cook, 2018), which is particularly suitable for identifying the major components of complex signals. LASSO regression across a broader time range initially identified major signal paths, followed by secondary optimization with 5-ns fine resolution around these critical paths, enhancing signal quality. This two-step processing strategy not only improved time-domain resolution but also preserved the most critical multipath structural features in the signal, outputting high-resolution time-domain impulse responses that more accurately reflect subtle signal changes caused by human activities.

Phase correction and format conversion constituted the final critical component of preprocessing. We processed the frequency domain signal matrix obtained from the previous LASSO optimization, converting it to standard format CSI data. The unwrap operation eliminated the phase jumps, and the subsequent two-stage phase correction further optimized the signal quality. The first stage detected and corrected sudden phase changes (phase differences exceeding 0.9π), while the second stage removed residual phase errors through linear fitting. This method effectively eliminates the system phase noise caused by internal clock drift and sampling bias in WiFi devices, preserves the phase information generated only by channel variations, and ultimately outputs standard format CSI data containing corrected amplitude and phase, which is suitable for subsequent sensing applications.

Through the above processing flow, the raw data are transformed into high-quality CSI matrices containing corrected amplitude and phase information. These optimized data can more accurately reflect the impact of human activities on wireless signals in the construction site environment, providing a foundation for subsequent feature extraction and activity recognition.

3.3.2. *Doppler spectral feature extraction*

After data preprocessing, it is necessary to extract features that accurately characterize human activities. The Doppler spectrum is particularly suitable, as it intuitively reflects velocity-related characteristics during human movement. This section describes the conversion process from the processed CFR matrices to Doppler spectrograms.

The preprocessed data had a shape of (500,58,2), where 500 represents the number of packets in the time sequence, 58 represents the number of subcarriers, and 2 represents amplitude and phase channels. These data packets contained rich temporal information, recording wireless signal change patterns caused by different human activities at construction site environments.

To effectively capture complete temporal features of activities, we employed a sliding window strategy to calculate Doppler spectrograms. Each window contained 140 consecutive data packets with a sliding step size of 2. The larger window size ensured that a complete action cycle could be captured within a single window while providing sufficient temporal resolution to record subtle changes during the action process. The smaller sliding step ensured fine-grained sampling in time, helping to capture rapidly changing movement details. The data shape of a single sliding window is (140,58,2), and a total of 180 windows are generated for the entire sliding process.

For data within each sliding window, we first applied a Hanning window function (Lessard, 2006) to reduce spectral leakage effects. This step effectively reduced pseudo-peaks caused by window truncation in FFT transformations, improving the accuracy of spectral estimation. Subsequently, we performed zero-padded FFT processing on the temporal signals within the window, appending 40 zeros to the original 140-point signal, expanding it to 180 points, and then executing a 180-point FFT transformation. This zero-padding operation enhanced spectral interpolation resolution, making velocity resolution more

refined.

The primary purpose of applying the FFT was to transform time-domain CSI amplitude variations into the frequency domain, thereby enabling the extraction of Doppler frequency shift features induced by human movement. According to the Doppler effect, the frequency shift of a wireless signal caused by a moving target is proportional to the target's radial velocity relative to the signal source. Through the FFT transform, these velocity-dependent frequency components are made explicit to distinguish the unique velocity features of different movement patterns. The Doppler vector output from the FFT for each window had a shape of (180,1), with each element in the vector corresponding to a specific velocity component, its value representing the energy intensity of that velocity component.

Finally, all 180-dimensional Doppler vectors produced by sliding windows were stacked in chronological order, forming a two-dimensional Doppler time-frequency spectrogram with a shape of (180,180). In this spectrogram, the horizontal axis represents time progression, determined by sliding window positions; the vertical axis represents Doppler frequency shifts, reflecting the target's radial movement velocity relative to the transmitter; and the pixel brightness represents the signal energy intensity at that time-frequency point.

This Doppler time-frequency spectral representation method intuitively displays unique time-varying frequency feature patterns produced by human activities. Compared to conventional time-domain or frequency-domain features, Doppler time-frequency spectrograms simultaneously contain information from both time and velocity dimensions, more comprehensively describing the dynamic characteristics of complex human activities. Especially for common movements at construction sites such as squatting, carrying bricks, climbing ladders, climbing stairs, and falling, Doppler spectrograms can capture the unique velocity change patterns of these movements, providing high-resolution, physically interpretable feature inputs for subsequent activity recognition.

3.4. Model architecture

The Doppler spectrograms obtained from previous steps contain rich time-frequency features that require extraction and classification through an effective deep learning architecture (Huizing et al., 2019; Kim and Moon, 2016). It has been shown that the lightweight Inception module proposed in SHARP (Meneghello et al., 2023) exhibits excellent classification performance in processing Doppler features, which stems from its unique multi-scale feature extraction capability that can effectively capture the multi-level changes in Doppler spectrograms generated by human movements. The time-frequency patterns embedded in Doppler spectrograms often have multi-scale structural characteristics, and the velocity changes of human activities have both macroscopic overall trends and microscopic fluctuations. The lightweight Inception architecture employs parallel multi-scale convolutional paths, enabling simultaneous extraction of features at different scales without a significant increase in computational complexity. This characteristic is particularly valuable for resource-constrained practical applications scenarios like construction sites. Based on these advantages, our research further optimized the Inception architecture to meet the specific requirements of fall detection tasks in construction site settings.

Our proposed CSI-based lightweight Inception-reduction fall detection network architecture consists of five core components, illustrated in Fig. 5. The input layer receives 180 × 180-dimensional Doppler spectrogram matrices, with the time dimension (180) representing consecutive Doppler vector sequences and the frequency dimension (180) corresponding to velocity resolution intervals, completely capturing time-varying velocity features of human movement.

The improved Inception-reduction module employs a four-path parallel structure for multi-scale feature extraction. The 2 × 2 maximum pooling branch with a stride of 2 captures the prominent feature representations in the spectrogram and effectively extracts the peak velocity features during the action. The 2 × 2 convolutional branch with a stride of 2 extracts mesoscopic-scale feature information, focusing on the overall structure of the action. The multi-scale convolution cascade branch implements gradual extraction of fine-grained features through a concatenated structure of 1 × 1, 2 × 2, and 4 × 4 convolution kernels, capturing minute but crucial movement details. And the 3 × 3 maximum pooling branch with a stride of 2 captures broader contextual information, enhancing global semantic understanding of actions. The feature maps from these four branches are concatenated along the channel dimension, achieving effective fusion of multi-scale features and forming a comprehensive characterization of motion.

To reduce model size while preserving key information, the feature optimization layer first employs a 1 × 1 convolution (with 4 kernels) to compress the originally high-dimensional channels to 4, preserving essential information while reducing computational cost. This dimensionality reduction process preserves the essential part of the information by learning linear combinations between features, while significantly reducing the number of parameters and computational burden. The reduced feature maps are flattened into one-dimensional vectors to provide input for fully connected layers. This design effectively reduces model complexity while preserving key feature information, improving training efficiency, and enhancing generalization capability.

Dropout regularization is integrated into the network, which temporarily deactivates neuron connections with a probability of 0.2 during the training phase, forming different network substructures to participate in each weight update. Subsequently, a fully connected layer

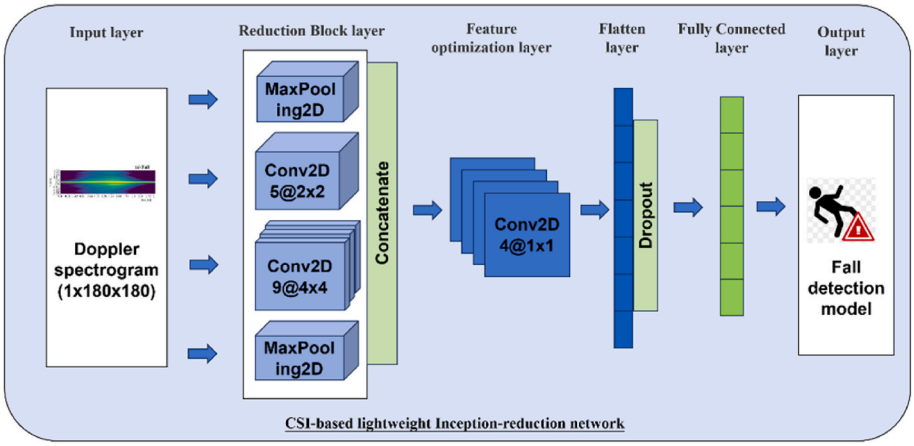

Fig. 5. The framework of the model.

takes the one-dimensional feature vector as input and performs linear combinations with the weights of each neuron. And the output layer produces binary classification results that clearly distinguish between fall and N-fall states. This output design meets the practical needs of safety monitoring at construction sites and can provide clear alert signals directly to the safety management system.

3.5. Evaluation metrics

To comprehensively evaluate the performance of the proposed method, this study employs multiple common classification evaluation metrics, including Precision, Recall, Accuracy, and F1 score. Precision reflects the proportion of actual falls among samples identified as falls, measuring the model's ability to reduce false positives; Recall represents the proportion of actual fall events correctly identified, measuring the model's ability to capture all real fall events; Accuracy provides a measure of overall classification performance; F1 score considers both precision and recall, offering a more balanced performance assessment. Given the application scenario of construction site fall detection systems, Recall holds particular importance in this study. In safety monitoring systems, missing fall events (false negatives) can lead to serious consequences, including delayed rescue and increased risk of injury. Therefore, while all metrics have reference values, improving recall is a priority consideration when designing the system.

4. Result

4.1. Analysis of features in spectrograms

The Doppler spectrograms visualize the time-frequency characteristics of different human activities and provide a physical perspective for understanding how CSI signals capture fall events at construction scenarios. Fig. 6 illustrates Doppler spectrograms of six typical construction activities, including falling, static environment, brick lifting, ladder climbing, squatting, and stair ascent/descent.

Falling motion (Fig. 6a) exhibits unique time-frequency pattern characteristics. Initially, from 0 to approximately 1.89 s, the spectrogram shows energy tightly concentrated around the 0 m/s axis, indicating the subject is in a stable, stationary state. Beginning at 1.89 s and intensifying until around 3.15 s, a dramatic event unfolds as the spectral energy rapidly disperses across a wide velocity range, peaking at approximately ±4.23 m/s. This creates a prominent vertical expansion in the spectrogram, corresponding to the critical phase where the individual loses balance and accelerates towards the ground. In the subsequent interval, from 3.15 to 4.41 s, the energy distribution begins to contract, but significant spectral fluctuations persist, reflecting the complex dynamics of the body impacting the ground and the subsequent, uncoordinated settling movements. These asynchronous limb movements appear as highly dispersed energy distributions in the spectrogram. After 4.41 s, the energy swiftly reconverges to the 0 m/s baseline and remains stable, signifying that the fall is complete and the subject has returned to a stationary state on the ground. The characteristic pattern of the fall is its rapid energy dispersion within a short period and its following quick convergence, which starkly contrasts with the other activities.

Static environment (Fig. 6b) serves as a reference baseline, displaying spectral characteristics when no human activity is present. Throughout the observation period, energy remains stably concentrated near 0 m/s, forming a narrow band. This feature reflects the scattering characteristics of static objects in the environment, with no apparent Doppler frequency shift phenomena.

Brick lifting activity (Fig. 6c) demonstrates spectral characteristics different from falling. During the 0.76–1.13 s interval, the spectral energy spreads towards ±1.41 m/s, reflecting the participant's start of the bending and reaching action, with the upper body tilting down to produce a negative velocity component, and the arm stretching downward also contributing to the negative motion signature. The energy spread reaches a maximum range (±2.82 m/s) between 1.13 and 1.89 s, corresponding to the critical phase when the participant grasps the brick and begins to lift it, with the legs, waist, and arms working in coordination to produce a complex pattern of movement. Between 1.89 and 2.27 s, spectral energy fluctuations gradually decrease, indicating the participant adjusting posture, possibly raising the brick to the target height. From 2.27 s until the end of the recording, energy primarily concentrates near 0 m/s, suggesting the participant has completed the lifting action and maintains a stable posture, possibly holding the brick with arms extended forward or placing it in a specific position. The spectral characteristic of brick lifting lies in its energy dispersion being less pronounced than falling, with changes occurring more gradually and in an orderly manner, reflecting a controlled, purposeful action.

Ladder climbing activity (Fig. 6d) presents distinctive periodic spectral characteristics. From 1.89 s onward, the spectrogram displays evident periodic alternating energy patterns, particularly during the active climbing phase between 1.89 and 7.18 s. This periodicity corresponds to the alternating hand and foot movement patterns during climbing: lifting legs generates positive velocity components, while arms grasping the next rung produce negative velocity components. Each periodic fluctuation in the spectrogram (approximately 1–1.5 s intervals) likely represents a complete climbing gait cycle, including leg lifting, grasping, body elevation, and other coordinated movements. The spectral characteristics of ladder climbing were mainly characterized by sustained periodic fluctuations rather than a single intense dispersion process seen in falling, reflecting a rhythmic, repetitive activity.

Squatting activity (Fig. 6e) exhibits distinct three-phase characteristics. During the 0–3.78 s preparation phase, spectral energy concentrates in a narrow band near 0 m/s, indicating the participant is in a relatively static standing posture, maintaining stability, preparing to begin the squatting motion. In the 3.78–5.67 s interval, spectral energy rapidly disperses to a range of ±2.82 m/s, forming a notable vertical expansion band. This phase corresponds to the complete squat-rise action sequence: the participant's center of gravity first moves downward, then recovers upward, and this bidirectional vertical movement appears as energy dispersion and fluctuation in the spectrogram. The key distinction between squatting and falling lies in the controllability of movement and symmetry of energy changes: squatting displays an orderly descent-ascent process, while falling is an uncontrolled unidirectional energy dispersion.

Stair ascent/descent activity (Fig. 6f) exhibits continuous periodic energy distribution concentrated within the ±2.12 m/s range. Throughout the 7.18-s observation period, the spectrogram displays stable alternating energy distribution patterns, reflecting the alternating elevation patterns of the participant's legs and torso during stair

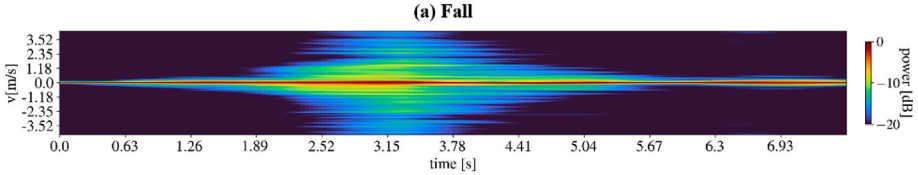

Fig. 6(A). Doppler spectrogram of falling motion.

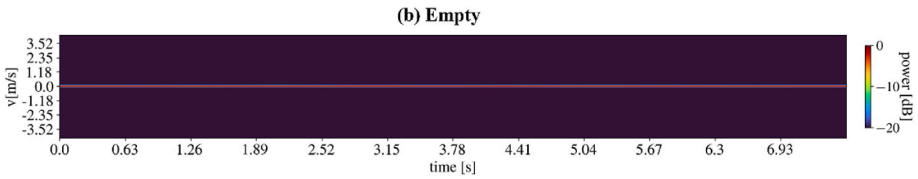

Fig. 6(B). Doppler spectrogram of a static environment.

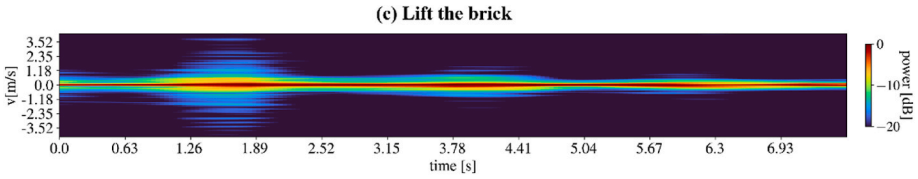

Fig. 6(C). Doppler spectrogram of brick lifting.

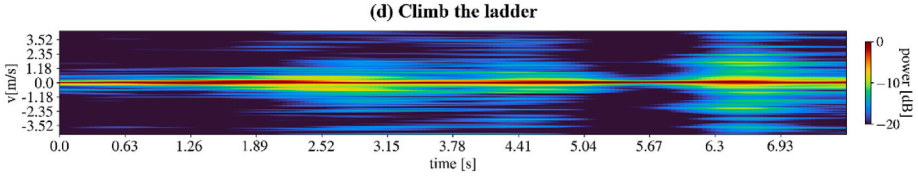

Fig. 6(D). Doppler spectrogram of ladder climbing.

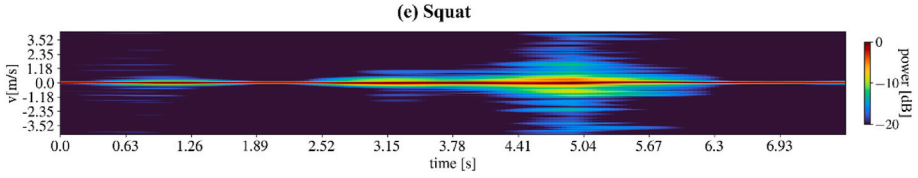

Fig. 6(e). Doppler spectrogram of squatting.

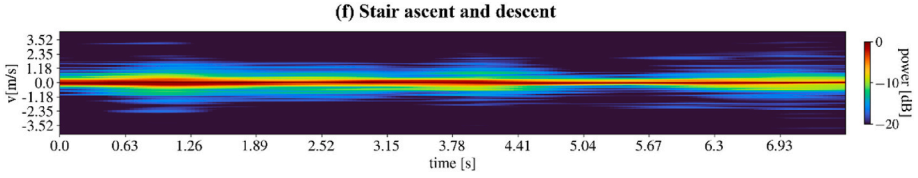

Fig. 6(f). Doppler spectrogram of stair ascent/descent.

movement. Compared to falling, the spectral changes in stair climbing are more uniform and sustained, with relatively stable energy distribution ranges, reflecting a regular, predictable human movement pattern.

Notably, the multipath characteristics of WiFi signal propagation significantly influence Doppler spectrograms. Even single-direction human movements may simultaneously present positive and negative frequency shifts in the spectrogram. This phenomenon can be understood through the following example, when a worker walks forward, relative to a WiFi receiver positioned in front of the worker, the body's movement direction forms an angle close to 0° with the signal propagation path, generating a positive Doppler frequency shift. At the same time, relative to the signal path reflected by the wall behind the worker, the body is moving away from this reflection path at an angle close to 180°, thus generating a negative Doppler shift. Furthermore, even within the same action, such as bending during brick lifting, different body parts generate different shift effects: the upper body bending downward produces negative frequency shifts, while the legs might generate slight positive shifts to maintain balance. These physical mechanisms collectively explain why both positive and negative frequency shifts appear simultaneously in spectrograms, especially pronounced in complex construction environments.

The Doppler spectral analysis results demonstrated that fall events exhibit distinct characteristics compared to other construction activities. Specifically, falls are marked by a short period of intense energy dispersion (within the range of ±4.23 m/s), followed by rapid spectral convergence. In contrast, other activities display their own unique but comparatively constrained spectral patterns. These observations provide a robust physical foundation for developing fall detection algorithms based on Doppler features.

4.2. *Analysis of overall model detection performance*

After analyzing the time-frequency domain features of different activity types, this section evaluates the overall performance of the CSI fall

detection model for mobile receivers constructed based on these features. To fully leverage the collected diversity, all 720 datasets, each representing a self-contained activity instance, were pooled and randomly shuffled before being partitioned into training, validation, and testing sets with a 6:2:2 ratio. This approach is consistent with methodologies in the WiFi sensing literature (Chu et al., 2023; Lee et al., 2020c), where similar random data partitioning is employed to validate generalized human activity recognition systems. Table 2 shows the confusion matrix of the model on the test set, where "Fall" represents fall events and "N-Fall" represents all non-fall activities (including common activities at construction sites such as lifting bricks, climbing ladders, squatting, and stair ascent/descent).

From the confusion matrix, the model correctly identified 63 out of 67 actual fall samples and only 4 were misclassified as non-falls; meanwhile, 72 out of 77 non-fall samples were correctly identified and 5 were misclassified as falls. This indicates that the model exhibits high accuracy in recognizing fall events and is particularly effective in reducing underreporting, which is crucial for construction site safety monitoring.

The performance metrics in Table 3 further confirm the effectiveness of the proposed model. The average accuracy across the two categories reaches 93.7 %, with balanced performance among all metrics, demonstrating the superiority of the Doppler effect-based feature extraction method. The slightly higher accuracy in the non-fall category indicates that the model effectively reduces false alarms, thereby enhancing the system's reliability and minimizing resource waste in practical scenarios. Notably, the fall category achieves a high recall of 94.0 %, indicating that most actual fall events are detected, which is critical for safety management on construction sites. Moreover, F1 scores exceeding 93 % for both categories confirm that the model achieves a favorable balance between precision and recall, providing a robust foundation for real-world deployment.

4.3. Robustness analysis of fall detection performance

To evaluate the practicality of the proposed system in complex construction environments, this section analyzes the model's performance under different distances. Table 4 displays detection results at three test distances of 2m, 6m, and 10m.

The model exhibits satisfactory stability under different distance conditions, as shown in Fig. 7. As distance increased, the recall maintained high levels above 90 % across all distance conditions. This relatively limited impact of distance variation on performance may be related to WiFi signal propagation characteristics and our system design. On the one hand, 2.4 GHz WiFi signals attenuate relatively slowly in open spaces, maintaining sufficient signal strength even at 8m distances. On the other hand, in this study, the receiver is directly worn on the target subject, ensuring that the signal receiving point maintains a fixed relative positional relationship with human movements. This mobile receiver configuration introduces beneficial multipath effects as the human body moves. When the receiver moves with the human body, it creates dynamic multipath propagation patterns that enrich the CSI features. These multipath effects can partially compensate for distance-related signal attenuation by providing additional signal diversity, ensuring that even as distance increases, sufficient information remains available for accurate activity detection. Regardless of how the distance between the human body and the transmitter changes, the receiving device can capture human motion features of similar quality. The precision for non-fall categories remained above 97 % at all distances, demonstrating excellent stability. Although F1 scores fluctuated slightly with distance changes, they maintained good levels above 85 %, verifying the model's reliability under different distance conditions.

Table 2
Confusion matrix of Fall.

	Predicted	
Actual	Fall	N-Fall
Fall	63	4
N-Fall	5	72

Table 3
Performance of the model in monitoring Fall.

	Precision	Recall	F1	Accuracy
Fall	92.7	94.0	93.3	93.7
N-Fall	94.7	93.5	94.1	93.8

Table 4
Performances of the model in monitoring Fall at different distances.

		Precision	Recall	F1	Accuracy
2m	Fall	81.5	91.7	86.2	93.1
	N-Fall	97.3	93.5	95.4	93.1
6m	Fall	77.2	94.4	85	93.7
	N-Fall	98.6	93.5	96	93.7
10m	Fall	82.8	96	88.9	94.1
	N-Fall	98.6	93.5	96	94.1

5. Discussion

5.1. Performance and contribution

This study introduces a fundamental shift in how WiFi sensing is applied to construction safety. Previous explorations in this domain have predominantly relied on static systems with both fixed transmitters and receivers. This configuration is inherently restrictive for dynamic construction environments, where work zones are constantly reconfigured and workers are highly mobile, making fixed-device solutions fundamentally difficult to scale and impractical for site-wide coverage. Our primary methodological innovation is the adoption of a mobile receiver worn by the worker. This transforms the sensing paradigm from monitoring a fixed space to creating a person-centric, dynamic sensing field. The core principle is that the receiver moves with the subject, ensuring that motion-induced signal variations are consistently and robustly captured regardless of the worker's absolute location relative to the transmitter.

The feasibility and effectiveness of this mobile-receiver framework were validated through a series of simulated experiments. The system demonstrated strong performance, achieving an overall accuracy of 93.7 %. Crucially for a safety application, it attained a high recall of 94.0 % for fall events, indicating a very low risk of missing detections. Furthermore, the system exhibited significant spatial robustness. As validated by our results, it maintained high accuracy (consistently above 93 %) and fall recall (above 90 %) across tested distances of 2m, 6m, and 10m. These findings confirm that our approach is not only innovative in principle but also effective and reliable in practice, offering a feasible and robust solution for fall detection at construction scenarios.

An interesting finding from our robust analysis is the system's stable, and even slightly enhanced, performance at greater distances, a seemingly counterintuitive result. This phenomenon indicates that performance in CSI sensing is governed by a complex interplay of factors, not just signal strength. Foundational work in this field has established that CSI provides fine-grained channel measurements across multiple subcarriers, capturing the detailed multipath propagation effects that are sensitive to human motion (Yang et al., 2013). At very close range, the dominant LoS signal may cause the system to be overly sensitive to minor, unintentional movements of the user holding the mobile receiver. At a moderate distance (6–10m), however, a more favorable sensing environment emerges. Here, the attenuated LoS signal becomes more balanced with multipath components, creating a richer

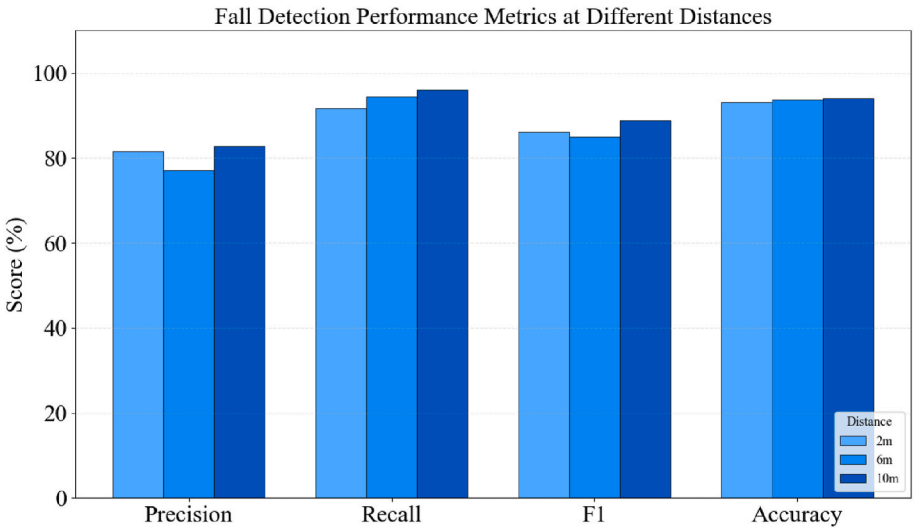

Fig. 7. Performance metrics at different distances.

interference pattern that provides a more distinctive signature for a fall event. Simultaneously, the relative impact of micro-movements from the user is diminished. This dual effect of increased multipath richness (Zhang et al., 2017) and reduced sensitivity to minor instabilities explains the system's robust performance.

5.2. Practical implications

The successful adoption of our method in a real-world construction site hinges on practical strategies for managing false alarms and integrating the system into safety workflows. A key mitigation strategy to enhance robustness is the strategic deployment of additional WiFi transmitters. Placing three to four transmitters around a high-activity work zone would create superior spatial diversity. This enriched signal environment ensures that a worker's movement is captured from multiple perspectives simultaneously, making the detection model less susceptible to signal blockages or dead spots from a single angle and inherently reducing the likelihood of false positives. This solution can be deployed effectively in sparsely populated work areas, overhead work zones, or other locations where constant manual supervision is difficult. Beyond the critical response required for individual falls, the system's value is magnified through the aggregation and analysis of incident data. A high frequency of fall events associated with a particular worker could serve as a non-punitive indicator of underlying risk factors, such as fatigue or health issues, enabling supportive intervention. Concurrently, mapping the locations of these incidents can reveal environmental hazards. A cluster of falls in a specific area, for instance, provides a clear signal to safety managers to inspect and remediate the location, such as addressing a slippery surface or an obstacle, thereby preventing future incidents for the entire workforce.

For practical implementation, the system needs to evolve from an offline, post-event analysis model to one capable of continuous, real-time monitoring. This would entail the workers' device maintaining a temporal buffer that continuously stores the most recent segment of CSI data (e.g., the last 5 s). A processing module, running either on the device or an edge server, would then analyze this data stream using a sliding window approach. By continuously processing overlapping windows of the most recent data, the system can ensure that any fall event is detected, and an alert is triggered within a few seconds of its occurrence. This design for a continuous analysis pipeline directly addresses the need for immediate responsiveness, making it a feasible on-site incident management tool.

5.3. Limitations and future directions

While this study demonstrates the possibility of the proposed approach for identifying falls among construction workers in a confined experimental setting, some limitations should be considered. Current experiments were conducted in a small-scale simulated construction site, and future validation in larger and more diverse construction environments is necessary. We acknowledge that the dataset, while diverse in activities and conditions, was collected from a limited number of participants. Future work should involve a larger and more varied cohort to further validate the model's generalizability across a wider population. Moreover, the smartphone used was the Nexus 6P, supporting Nexmon firmware. Requiring all workers to use a specific smartphone model is impractical in real applications. Subsequent research should consider hardware device diversity, evaluate different receiving devices' impact on accuracy, or develop more universal signal reception solutions. The optimal placement strategy for WiFi routers at construction sites represents another worthwhile research direction. Signal propagation at construction environments is influenced by site layout, construction materials, and obstacles. Determining optimal router placement can significantly improve fall detection accuracy and effectiveness. In the future, we plan to integrate this technology with Building Information Modeling and site management systems to achieve a more comprehensive safety monitoring solution.

Furthermore, while our experiments included ambient interference from nearby traffic and pedestrians, we did not systematically investigate the effect of high-intensity interference occurring in very close proximity to the monitored worker. For instance, the scenario where another worker passes directly between the subject and the transmitter, or where a large metallic object like a tool cart is moved nearby during a fall event, was not specifically addressed. The Doppler signatures from such close and significant movements could potentially impact the model's performance. Therefore, future research should focus on collecting a more diverse dataset that includes these specific close-proximity interference events to further evaluate and enhance the model's robustness for deployment in crowded and highly dynamic site conditions.

6. Conclusion

This study introduced and validated a new framework for fall detection among construction workers, leveraging mobile WiFi-CSI signals captured by smartphones. The principal contribution lies in shifting the paradigm from static, area-based monitoring to a person-

centric approach, thereby addressing the inherent limitations of conventional methods regarding installation cost, privacy concerns, and adaptability in dynamic environments. A relationship between several common construction activities and fall events, and CSI Doppler features is established by validating the Doppler frequency shift characteristics of CSI signals under the utilization of receiver movement.

The empirical results demonstrated the system's high efficacy and spatial robustness. It achieved overall accuracy exceeding 93 % across various distances and consistently maintained F1-scores above 85 % for fall events, confirming its reliability for practical deployment. It provides a more flexible and practical solution for construction site safety monitoring, enabling project managers to achieve monitoring of workers through smartphones carried with them, and provide timely intervention and necessary medical support when a potential fall is detected, thus significantly improving the efficiency of site safety management. With further development, the system is expected to be integrated with intelligent site management systems to form a comprehensive safety monitoring network.

CRediT authorship contribution statement

Yinong Hu: Writing – review & editing, Writing – original draft, Visualization, Validation, Investigation, Formal analysis, Data curation, Conceptualization. **Heng Li:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Mingzhou Cheng:** Writing – review & editing, Visualization, Validation, Supervision, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Mingyu Zhang:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Shuai Han:** Writing – review & editing, Validation, Methodology, Conceptualization. **Waleed Umer:** Writing – review & editing.

Declaration of competing interest

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 paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China Grant (42302322).

Data availability

Data will be made available on request.

References

Abuhoureyah, F.S., Wong, Y.C., Mohd Isira, A.S.B., 2024. WiFi-based human activity recognition through wall using deep learning. Eng. Appl. Artif. Intell. 127, 107171. https://doi.org/10.1016/j.engappai.2023.107171.

Al-qaness, M.A.A., Li, F., Ma, X., Zhang, Y., Liu, G., 2016. Device-free indoor activity recognition system. Appl. Sci. 6, 329. https://doi.org/10.3390/app6110329.

Alwan, M., Rajendran, P.J., Kell, S., Mack, D., Dalal, S., Wolfe, M., Felder, R., 2006. A smart and passive floor-vibration based fall detector for elderly. In: 2006 2nd Int. Conf. Inf. Commun. Technol., pp. 1003–1007. https://doi.org/10.1109/ICTTA.2006.1684511.

Antwi-Afari, M.F., Li, H., Seo, J., Wong, A.Y.L., 2018. Automated detection and classification of construction workers' loss of balance events using wearable insole pressure sensors. Autom. ConStruct. 96, 189–199. https://doi.org/10.1016/j.autcon.2018.09.010.

Antwi-Afari, M.F., Li, H., Umer, W., Yu, Y., Xing, X., 2020. Construction activity recognition and ergonomic risk assessment using a wearable insole pressure system. J. Construct. Eng. Manag. 146, 04020077. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001849.

Bulletin2022_issue23_sc. https://www.labour.gov.hk/common/osh/pdf/Bulletin2022_issue23_sc.pdf, 2022–. (Accessed 12 October 2023).

Chaccour, K., Darazi, R., El Hassani, A.H., Andrès, E., 2017. From fall detection to fall prevention: a generic classification of fall-related systems. IEEE Sens. J. 17, 812–822. https://doi.org/10.1109/JSEN.2016.2628099.

Choo, H., Lee, B., Kim, H., Choi, B., 2023a. Automated detection of construction work at heights and deployment of safety hooks using IMU with a barometer, autom. Constr. Met. (CTICM) 147, 104714. https://doi.org/10.1016/j.autcon.2022.104714.

Choo, H., Lee, B., Kim, H., Choi, B., 2023b. Automated detection of construction work at heights and deployment of safety hooks using IMU with a barometer, autom. Constr. Met. (CTICM) 147, 104714. https://doi.org/10.1016/j.autcon.2022.104714.

Chu, Y., Cumanan, K., Sankarpandi, S.K., Smith, S., Dobre, O.A., 2023. Deep learning-based fall detection using WiFi channel State information. IEEE Access 11, 83763–83780. https://doi.org/10.1109/ACCESS.2023.3300726.

Courtney, T.K., Sorock, G.S., Manning, D.P., Collins, J.W., Holbein-Jenny, M.A., 2001. Occupational slip, trip, and fall-related injuries can the contribution of slipperiness be isolated? Ergonomics 44, 1118–1137. https://doi.org/10.1080/00140130110085538.

Ding, L., Fang, W., Luo, H., Love, P.E.D., Zhong, B., Ouyang, X., 2018. A deep hybrid learning model to detect unsafe behavior: integrating convolution neural networks and long short-term memory. Autom. ConStruct. 86, 118–124. https://doi.org/10.1016/j.autcon.2017.11.002.

Duan, P., Goh, Y.M., Zhou, J., 2023a. Personalized stability monitoring based on body postures of construction workers working at heights. Saf. Sci. 162, 106104. https://doi.org/10.1016/j.ssci.2023.106104.

Duan, P., Goh, Y.M., Zhou, J., 2023b. Personalized stability monitoring based on body postures of construction workers working at heights. Saf. Sci. 162, 106104. https://doi.org/10.1016/j.ssci.2023.106104.

Dzeng, R.-J., Fang, Y.-C., Chen, I.-C., 2014. A feasibility study of using smartphone built-in accelerometers to detect fall portents. Autom. ConStruct. 38, 74–86. https://doi.org/10.1016/j.autcon.2013.11.004.

Fang, Y.-C., Dzeng, R.-J., 2014. A smartphone-based detection of fall portents for construction workers. Procedia Eng. 85, 147–156. https://doi.org/10.1016/j.proeng.2014.10.539.

Fang, W., Zhong, B., Zhao, N., Love, P.E.D., Luo, H., Xue, J., Xu, S., 2019a. A deep learning-based approach for mitigating falls from height with computer vision: convolutional neural network. Adv. Eng. Inform. 39, 170–177. https://doi.org/10.1016/j.aei.2018.12.005.

Fang, W., Zhong, B., Zhao, N., Love, P.E.D., Luo, H., Xue, J., Xu, S., 2019b. A deep learning-based approach for mitigating falls from height with computer vision: convolutional neural network. Adv. Eng. Inform. 39, 170–177. https://doi.org/10.1016/j.aei.2018.12.005.

Fang, W., Love, P.E.D., Luo, H., Ding, L., 2020. Computer vision for behaviour-based safety in construction: a review and future directions. Adv. Eng. Inform. 43, 100980. https://doi.org/10.1016/j.aei.2019.100980.

Ferreira De Sousa, F.A.S., Escriba, C., Avina Bravo, E.G., Brossa, V., Fourniols, J.-Y., Rossi, C., 2022. Wearable pre-impact fall detection system based on 3D accelerometer and subject's height. IEEE Sens. J. 22, 1738–1745. https://doi.org/10.1109/JSEN.2021.3131037.

Fugate, H., Alzraiee, H., 2023. Quantitative analysis of construction labor acceptance of wearable sensing devices to enhance workers' safety. Results Eng. 17, 100841. https://doi.org/10.1016/j.rineng.2022.100841.

Guan, Q., Li, C., Guo, X., Shen, B., 2017. Infrared signal based elderly fall detection for In-Home monitoring. In: 2017 9th Int. Conf. Intell. Hum.-Mach. Syst. Cybern. IHMSC, pp. 373–376. https://doi.org/10.1109/IHMSC.2017.91.

Guo, R., Li, H., Han, D., Liu, R., 2023. Feasibility analysis of using channel state information (CSI) acquired from Wi-Fi routers for construction worker fall detection. Int. J. Environ. Res. Publ. Health 20, 4998. https://doi.org/10.3390/ijerph20064998.

Halperin, D., Hu, W., Sheth, A., Wetherall, D., 2011. Tool release: gathering 802.11n traces with channel state information. ACM SIGCOMM 41, 53. https://doi.org/10.1145/1925861.1925870.

Hong, S., Yoon, J., Ham, Y., Lee, B., Kim, H., 2023. Monitoring safety behaviors of scaffolding workers using Gramian angular field convolution neural network based on IMU sensing data. Autom. ConStruct. 148, 104748. https://doi.org/10.1016/j.autcon.2023.104748.

Huizing, A., Heiligers, M., Dekker, B., de Wit, J., Cifola, L., Harmanny, R., 2019. Deep learning for classification of Mini-UAVs using micro-doppler spectrograms in cognitive radar. IEEE Aero. Electron. Syst. Mag. 34, 46–56. https://doi.org/10.1109/MAES.2019.2933972.

Kim, Y., Moon, T., 2016. Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks. IEEE Geosci. Remote Sens. Lett. 13, 8–12. https://doi.org/10.1109/LGRS.2015.2491329.

Kim, Y., Jung, H., Koo, B., Kim, J., Kim, T., Nam, Y., 2020. Detection of pre-impact falls from heights using an inertial measurement unit sensor. Sensors 20, 5388. https://doi.org/10.3390/s20185388.

Lee, H., Ahn, C.R., Choi, N., 2020a. Fine-grained occupant activity monitoring with Wi-Fi channel state information: practical implementation of multiple receiver settings. Adv. Eng. Inform. 46, 101147. https://doi.org/10.1016/j.aei.2020.101147.

Lee, Y.-C., Shariatfar, M., Rashidi, A., Lee, H.W., 2020b. Evidence-driven sound detection for prenotification and identification of construction safety hazards and accidents. Autom. ConStruct. 113, 103127. https://doi.org/10.1016/j.autcon.2020.103127.

Lee, H., Ahn, C.R., Choi, N., 2020c. Fine-grained occupant activity monitoring with Wi-Fi channel state information: practical implementation of multiple receiver settings. Adv. Eng. Inform. 46, 101147. https://doi.org/10.1016/j.aei.2020.101147.

Lessard, C.S., 2006. Window functions and spectral leakage. In: Signal Process. Random Physiol. Signals. Springer International Publishing, Cham, pp. 175–193. https://doi.org/10.1007/978-3-031-01610-3_16.

Lin, Z., Wang, Z., Dai, H., Xia, X., 2022. Efficient fall detection in four directions based on smart insoles and RDAE-LSTM model. Expert Syst. Appl. 205, 117661. https://doi.org/10.1016/j.eswa.2022.117661.

Lipscomb, H.J., Glazner, J.E., Bondy, J., Guarini, K., Lezotte, D., 2006. Injuries from slips and trips in construction. Appl. Ergon. 37, 267–274. https://doi.org/10.1016/j.apergo.2005.07.008.

Liu, J., Zeng, Y., Gu, T., Wang, L., Zhang, D., 2021. WiPhone: smartphone-based respiration monitoring using ambient reflected WiFi signals. Proc ACM Interact Mob Wearable Ubiquitous Technol 5 (23), 1–23:19. https://doi.org/10.1145/3448092.

Liu, X., Xu, F., Zhang, Z., Sun, K., 2023. Fall-portent detection for construction sites based on computer vision and machine learning. Eng. Construct. Architect. Manag. 32, 1499–1521. https://doi.org/10.1108/ECAM-05-2023-0458.

M, A., Y, N.J., H, K.N., 2023a. Multimodal spatiotemporal skeletal kinematic gait feature fusion for vision-based fall detection. Expert Syst. Appl. 212, 118681. https://doi.org/10.1016/j.eswa.2022.118681.

M, A., Y, N.J., H, K.N., 2023b. Multimodal spatiotemporal skeletal kinematic gait feature fusion for vision-based fall detection. Expert Syst. Appl. 212, 118681. https://doi.org/10.1016/j.eswa.2022.118681.

Meneghello, F., Garlisi, D., Fabbro, N.D., Tinnirello, I., Rossi, M., 2023. SHARP: environment and person independent activity recognition with commodity IEEE 802.11 access points. IEEE Trans. Mobile Comput. 22, 6160–6175. https://doi.org/10.1109/TMC.2022.3185681.

Nyan, M.N., Tay, F.E.H., Tan, A.W.Y., Seah, K.H.W., 2006. Distinguishing fall activities from normal activities by angular rate characteristics and high-speed camera characterization. Med. Eng. Phys. 28, 842–849. https://doi.org/10.1016/j.medengphy.2005.11.008.

Occupational Safety and Health Management in the Construction Sector, 2021. International Labour Organization. https://www.ilo.org/meetings-and-events/occupational-safety-and-health-management-construction-sector-3.

OSH_Statistics_2022_tc. https://www.labour.gov.hk/common/osh/pdf/OSH_Statistics_2022_tc.pdf, 2022–. (Accessed 12 October 2023).

P, S., Udgata, S.K., 2024. Wi-Fi sensing based person identification and activity recognition using two-phase deep learning model. Eng. Appl. Artif. Intell. 132, 107904. https://doi.org/10.1016/j.engappai.2024.107904.

Patil, P., Patil, M.R., Itraj, S., Bomble, U.L., 2017. A review on MIMO OFDM technology basics and more. In: 2017 Int. Conf. Curr. Trends Comput. Electr. Electron. Commun. CTCEEC, pp. 119–124. https://doi.org/10.1109/CTCEEC.2017.8455114.

Poon, S.W., 2009. Workers' compensation for non-fatal construction accidents: review of Hong Kong court cases. Asian Soc. Sci. 5.

Qian, K., Wu, C., Yang, Z., Liu, Y., Jamieson, K., 2017. Widar: decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi. In: Proc. 18th ACM Int. Symp. Mob. Ad Hoc Netw. Comput. Association for Computing Machinery, New York, NY, USA, pp. 1–10. https://doi.org/10.1145/3084041.3084067.

Ranstam, J., Cook, J.A., 2018. LASSO regression. Br. J. Surg. 105, 1348. https://doi.org/10.1002/bjs.10895.

Santoyo-Ramón, J.A., Casilari, E., Cano-García, J.M., 2018. Analysis of a smartphone-based architecture with multiple mobility sensors for fall detection with supervised learning. Sensors 18, 1155. https://doi.org/10.3390/s18041155.

Sarkar, T.K., Ji, Z., Kim, K., Medouri, A., Salazar-Palma, M., 2003. A survey of various propagation models for mobile communication. IEEE Antenn. Propag. Mag. 45, 51–82. https://doi.org/10.1109/MAP.2003.1232163.

Sayeed, A.M., Aazhang, B., 1999. Joint multipath-Doppler diversity in mobile wireless communications. IEEE Trans. Commun. 47, 123–132. https://doi.org/10.1109/26.747819.

Schirn, A., 2024. OSHA Top 10 Violations of 2023 - ANSI Blog. ANSI Blog. https://blog.ansi.org/ansi/osha-top-10-violations-of-2023/.

Schulz, M., Wegemer, D., Hollick, M., 2016. DEMO: using NexMon, the C-based WiFi firmware modification framework. In: Proc. 9th ACM Conf. Secur. Priv. Wirel. Mob. Netw. Association for Computing Machinery, New York, NY, USA, pp. 213–215. https://doi.org/10.1145/2939918.2942419.

Seo, J., Han, S., Lee, S., Kim, H., 2015. Computer vision techniques for construction safety and health monitoring. Adv. Eng. Inform. 29, 239–251. https://doi.org/10.1016/j.aei.2015.02.001.

Shalaby, E., ElShennawy, N., Sarhan, A., 2022. Utilizing deep learning models in CSI-based human activity recognition. Neural Comput. Appl. 34, 5993–6010. https://doi.org/10.1007/s00521-021-06787-w.

Tan, B., Chen, Q., Chetty, K., Woodbridge, K., Li, W., Piechocki, R., 2018. Exploiting WiFi channel state information for residential healthcare informatics. IEEE Commun. Mag. 56, 130–137. https://doi.org/10.1109/MCOM.2018.1700064.

Tang, Z., Liu, Q., Wu, M., Chen, W., Huang, J., 2021. WiFi CSI gesture recognition based on parallel LSTM-FCN deep space-time neural network. China Commun 18, 205–215. https://doi.org/10.23919/JCC.2021.03.016.

Thariq Ahmed, H.F., Ahmad, H., Aravind, C.V., 2020. Device free human gesture recognition using Wi-Fi CSI: a survey. Eng. Appl. Artif. Intell. 87, 103281. https://doi.org/10.1016/j.engappai.2019.103281.

Wang, H., Ho, I.W.-H., 2024. CSI-based passenger counting on public transport vehicles with multiple transceivers. In: 2024 IEEE Wirel. Commun. Netw. Conf. WCNC, pp. 1–6. https://doi.org/10.1109/WCNC57260.2024.10571130.

Wang, Y., Wu, K., Ni, L.M., 2017a. WiFall: device-Free fall detection by wireless networks. IEEE Trans. Mobile Comput. 16, 581–594. https://doi.org/10.1109/TMC.2016.2557792.

Wang, H., Zhang, D., Wang, Y., Ma, J., Wang, Y., Li, S., 2017b. RT-Fall: a real-time and contactless fall detection System with commodity WiFi devices. IEEE Trans. Mobile Comput. 16, 511–526. https://doi.org/10.1109/TMC.2016.2557795.

Yan, B., Cheng, W., Li, Y., Gao, X., Liu, H., 2022. Joint activity recognition and indoor localization with WiFi sensing based on multi-view fusion strategy, digit. Signal Process. 129, 103680. https://doi.org/10.1016/j.dsp.2022.103680.

Yang, Z., Zhou, Z., Liu, Y., 2013. From RSSI to CSI: indoor localization via channel response. ACM 46, 25:1–25:32. https://doi.org/10.1145/2543581.2543592.

Yang, K., Jebelli, H., Ahn, C.R., Vuran, M.C., 2015. Threshold-Based Approach to Detect Near-Miss Falls of Iron Workers Using Inertial Measurement Units, pp. 148–155. https://doi.org/10.1061/9780784479247.019.

Yousefi, S., Narui, H., Dayal, S., Ermon, S., Valaee, S., 2017. A survey on behavior recognition using WiFi channel state information. IEEE Commun. Mag. 55, 98–104. https://doi.org/10.1109/MCOM.2017.1700082.

Zhang, D., Wang, H., Wu, D., 2017. Toward centimeter-scale human activity sensing with Wi-Fi signals. Computer 50, 48–57. https://doi.org/10.1109/MC.2017.7.