Check for updates

scientific reports



Yang Zhang¹, Xuechun Wang², Jinghao Wen³ & Xianxun Zhu⁴

In the swiftly evolving landscape of Internet of Things (IoT) technology, the demand for adaptive non-contact sensing has seen a considerable surge. Traditional human perception technologies, such as vision-based approaches, often grapple with problems including lack of sensor versatility and sub-optimal accuracy. To address these issues, this paper introduces a novel, non-contact method for human presence perception, relying on WiFi. This innovative approach involves a sequential process, beginning with the pre-processing of collected Channel State Information (CSI), followed by feature extraction, and finally, classification. By establishing signal models that correspond to varying states, this method enables the accurate perception and recognition of human presence. Remarkably, this technique exhibits a high level of precision, with sensing accuracy reaching up to 99%. The potential applications of this approach are extensive, proving to be particularly beneficial in contexts such as smart homes and healthcare, amongst various other everyday scenarios. This underscores the significant role this novel method could play in enhancing the sophistication and effectiveness of human presence detection and recognition systems in the IoT era.

Keywords Non-contact, Human presence sensing, Wireless perception, Machine learning

In contemporary society, the observation and interpretation of human activities carry substantial societal value. The choice of appropriate technologies for human perception has emerged as a significant area of modern research¹⁻³. Conventional human perception technologies mainly include vision⁴, sensors⁵, and infrared⁶. Visionbased methods entail collecting images or videos through a camera, followed by the application of image processing algorithms for recognition and perception. Despite its high accuracy, this technique is notably vulnerable to lighting conditions' variability, and the capture of video images may encroach on individual privacy. The sensor-based approach requires users to constantly wear sensors, potentially causing interruptions in their daily activities. Infrared-based methods, employing infrared sensors for human perception, are plagued by a high frequency of false alarms and are easily obstructed.

In 2011, aiming to surmount the limitations of traditional technologies, Halperin et al.⁷ developed a WiFi device firmware based on the IEEE 802.11n standard. This pioneering effort aimed to simplify the collection of CSI, laying the groundwork for human perception recognition. Subsequent to this advancement, various research groups and scholars have thoroughly explored CSI-based human perception technologies^{8,9,24,25}. In 2016, the WiFi ID method proposed by Zhang et al.¹⁰ utilized Fourier transformation and the relief algorithm to extract gait information, performing gait recognition via the Support Vector Machine (SVM) algorithm and achieving an average accuracy rate of 93%. In 2017, Shi et al.¹¹ executed human identification using a neural network model, attaining an accuracy rate of 94%. Despite the swift progression in the theoretical foundation and practical application of wireless sensing technology^{12,13}, significant challenges such as inadequate robustness, low accuracy, and limited universality, continue to prevail^{26–28}. This paper introduces a cutting-edge, non-contact human presence detection technology based on wireless sensing. The interpretation of the gathered sensing information allows for the identification of individuals within the sensing area, thereby enabling high-precision, non-contact sensing.

System block diagram

The block diagram of the WiFi-based non-contact human presence sensing system proposed in this paper is presented in Fig. 1. Initially, a network card with modified firmware is utilized to gather channel state information. Following this, noise is mitigated using a low-pass filter and wavelet transform. Subsequently, an algorithm founded on a neural network is employed to extract distinguishing features. Ultimately, machine learning techniques are used to classify and identify human states.

¹School of Economics and Management, Shanghai Polytechnic University, Shanghai 201209, China. ²School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan 430068, China. ³School of Computer Science, Central China Normal University, Wuhan 430079, China. ⁴School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China. email: zhangyang@sspu.edu.cn





Relevant theories

According to the IEEE 802.11 standard, the process of WiFi signal transmission is segmented into various subcarriers, each operating at a distinct frequency. The propagation paths of these subcarriers through the environment differ, leading to the diversity observed in the CSI data¹⁴. The CSI encompasses a range of data, including the time delay, amplitude attenuation, and phase shift experienced during the signal's transmission and reception. Essentially, CSI represents the physical layer information of the subcarriers, characterizing the cumulative effect on wireless signals at the receiver after undergoing reflection, refraction, and scattering across different environmental paths²⁹

In the frequency domain, the communication model for a system employing multiple transmitting and receiving antennas, characterized by narrowband flat fading, is represented as follows:

$$y = Hx + n \tag{1}$$

Here, y denotes the received signal, x symbolizes the transmitted signal, n represents noise, and H is the signal transformation matrix. This matrix H reflects the impact of the external physical environment on the transmitted signal x, as it propagates from the transmitter to the receiver, transforming into y. The matrix H can also be estimated using the following equation:

$$\hat{H} = \frac{\gamma}{x} \tag{2}$$

CSI is essentially a representation of H. At the receiver, the CSI for each subcarrier can be quantified in terms of amplitude and phase, as per the equation:

İ

$$H_i = \|H_i\| \exp(j \angle H_i) \tag{3}$$

In this notation, H_i signifies the value of the *i*th subcarrier in the Channel State Information, $||H_i||$ represents the amplitude of the *i*th subcarrier, and $\angle H_i$ denotes the phase of the *i*th subcarrier³⁰. The Signal-to-Noise Ratio (SNR) plays a critical role in this context, as it quantifies the level of the desired signal relative to the background noise, which is crucial in analyzing the quality and reliability of the communication channel. Variations in the environmental conditions during the wireless signal transmission can lead to multipath propagation, encompassing a line-of-sight path and several paths involving reflection and refraction^{15,22}. The crux of this paper is to analyze the alterations in channel propagation, induced by environmental changes during the propagation process, to facilitate the detection of human presence.

Methodology Pretreatment

The channel state information acquired directly is susceptible to low-pass noise and, therefore, cannot be directly employed for human presence detection. Accordingly, a low-pass filtering method is chosen²⁰. This technique exploits the characteristics of inductors with high resistance at low frequencies and capacitors with low resistance at high frequencies to perform data denoising. The formula is as follows:

$$f_c = \frac{1}{2\pi RC} \tag{4}$$

$$L(s) = \frac{2\pi f_c}{2\pi f_c + s} \tag{5}$$

where *R* is the resistance, *l* is the load series inductor, *C* is the parallel capacitor at both ends of the load resistance, and f_c is the cutoff frequency.

To augment the detection of indoor human activities, wavelet transforms¹⁷, recognized for their aptitude to differentiate high and low-frequency components, are employed for noise reduction. The core principle hinges on signal extraction through localized transformations in both spatial and frequency domains. By employing scaling and translation, the CSI is analyzed multi-scale. This approach distinctly segregates high and low-frequency components, enhancing the robustness and precision of the sensing technology. The mathematical representations of the hard($\phi_{\lambda hard}(\varpi)$) and soft ($\phi_{\lambda soft}(\varpi)$) thresholding functions are shown in Eqs. (6) and (7), respectively.

$$\phi_{\lambda \text{hard}}(\varpi) = \begin{cases} \varpi, \ | \ \varpi \ | \ge \lambda \\ 0, \ | \ \varpi \ | < \lambda \end{cases}$$
(6)

$$\phi_{\lambda \text{soft}}(\varpi) = \begin{cases} \text{sgn}(\varpi)(|\varpi| - \lambda), |\varpi| \ge \lambda \\ 0, & |\varpi| < \lambda \end{cases}$$
(7)

where ϖ is the wavelet coefficient, λ the threshold value, and sgn(.) denotes the signum function. In hard thresholding, the coefficient is nullified if its absolute value is below the threshold, while preserved otherwise. In contrast, the soft thresholding reduces the absolute value of each coefficient by λ and sets it to zero if the result is non-positive.

Feature extraction

While preprocessing significantly improves signal quality, the extensive volume of data complicates the direct classification and interpretation of human activities. A self-organizing neural network¹⁸, utilizing unsupervised learning with a competitive approach, skillfully extracts channel state information features following preprocessing. This network primarily consists of input and output layers, dedicated to classification and clustering tasks. The operation sequence is as follows:

The data first undergoes normalization to facilitate uniformity in subsequent processes. This step is mathematically represented as:

$$\hat{X} = \frac{X}{\|X\|} \tag{8}$$

Post-normalization, a similarity metric identifies the most relevant neurons. This involves the normalization of the weight vectors, expressed as:

$$\hat{W}_j = \frac{W_j}{\|W_j\|} \tag{9}$$

where \hat{X} and \hat{W}_i denote the normalized input and weight vectors, respectively.

The core of the competitive learning algorithm is to minimize the distance between the input vector and the weight vector of the selected neuron. This is encapsulated by the following equation:

$$\|\hat{X} - \hat{W_{j^*}}\| = \min_{j \in \{1, 2, 3, \dots, n\}} \{ \|\hat{X} - \hat{W_j}\| \}$$
(10)

where j^* represents the index of the winning neuron.

Subsequently, the output values are updated as follows:

$$y_j(t+1) = \begin{cases} 1, \text{ if } j = j^* \\ 0, \text{ if } j \neq j^* \end{cases}$$
(11)

where $y_i(t + 1)$ denotes the output of the *j*th neuron at time (t + 1).

For the winning neuron, the weight vector is refined using the equation:

$$\hat{W}_{j^*}(t+1) = \hat{W}_{j^*}(t) + \alpha \left(\hat{X} - \hat{W}_{j^*} \right)$$
(12)

Here, $\hat{W}_{j*}(t)$ is the weight vector at time *t*, adjusted to reduce the disparity with the input vector \hat{X} . The term α , representing the learning rate, falls within the range (0, 1]. It regulates the adaptation pace and magnitude of the weight vectors.

As the iteration count *t* progresses, the learning rate α gradually reduces to zero. This decline ensures the convergence and stability of the learning algorithm, thus preventing any potential overshooting of the optimal weight configuration. This methodical reduction in α is pivotal for the efficacy and efficiency of the learning process¹⁶.

Classification

In the pursuit of classifying and discerning features indicative of human presence, this study employs a softmax classifier¹⁹. This classifier adeptly computes probabilities for various states corresponding to different feature vectors. The operational flow of the softmax classifier is outlined as follows:

Initially, data is introduced into the input layer. It then traverses through two distinct feature layers, undergoing processing and transformation. Conclusively, the softmax function is applied, ensuring that each output is normalized to a probability range of 0 to 1. This normalization is formally represented as:

$$\operatorname{Softmax}(x)_{i} = \frac{e^{x_{i}}}{\sum_{i=1}^{K} e^{x_{j}}}$$
(13)

for an input vector $x \in \mathbb{R}^{K}$, where *i* is the index of a particular element, and *K* is the total number of classes.

Through this process, the classifier effectively transforms raw data into a probability distribution, facilitating the interpretation of each output as a conditional probability under various scenarios. This probabilistic framework allows for a more nuanced and accurate classification, pivotal in the intricate task of human presence feature recognition.

Ethics approval

This article does not contain any studies with human participants performed by any of the authors.

Experiment and analysis

Experimental setup

In response to the necessity for detailed experimental methodologies and architectural insights, this document meticulously delineates the experimental procedures, ensuring lucidity and reproducibility.

Experimental setup: The CSI dimension in our study is 1*3*30. This configuration comprises 1 transmitting antenna and 3 receiving antennas. The experimental hardware involves two Lenovo desktop computers, each powered by Intel Core i5800 CPUs. The transmitting computer is equipped with a single antenna, while the receiving computer features three antennas. The wireless transceiving system includes a Monitor Point (MP) for signal reception and an Access Point (AP) for signal transmission. The spatial arrangement of these antennas is illustrated in Fig. 2.

Software and model construction: The model is developed on MATLAB, a platform well-regarded for its robust capabilities in algorithm development and simulation. To ensure a consistent operational environment, both computers utilize Ubuntu 14.04 LTS. Equipped with Intel 5300 network cards, these systems are subjected to precise kernel and driver configurations prior to the installation of the CSI toolbox, as referenced in²³. This toolbox is instrumental in processing the CSI, offering a comprehensive suite of functionalities for analyzing and interpreting wireless signal characteristics. It facilitates the extraction and manipulation of CSI data, essential for our research. Moreover, the Self-Organizing Map neural network employed in our study is configured with a default output dimension of 100, optimizing the performance for our specific application.

Data collection environment and process: Data was gathered in two distinct environments to evaluate the model's versatility and efficacy under various conditions. The first environment was a laboratory measuring 6.5 m by 8 m, filled with test benches, chairs, and computers, presenting significant obstructions and strong multipath interference. The second was a spacious conference room, sized 9.5 m by 11 m, where multipath interference was minimal. In each setting, 200 datasets representing both occupied and unoccupied states were meticulously collected, with each session lasting 180 s and involving the transmission of 100 Channel State Information packets per second. The data was collected at different speeds and by different personnel, with a training and testing set ratio of 8:2.

Dataset and validation: Due to the lack of a standardized dataset for human presence detection, our research utilizes a proprietary dataset for model validation. This dataset covers two distinct scenarios, facilitating a thorough evaluation.

Evaluation metrics: To comprehensively evaluate the performance of our experiment, we utilized the True Positive Rate (TPR) and False Positive Rate (FPR) as key metrics. The True Positive Rate refers to the probability





Figure 2. Experimental environment plan.

of successfully detecting the presence of a person in the test set when someone is actually present. Conversely, the False Positive Rate indicates the probability of incorrectly identifying the presence of a person when, in reality, no one is present.

Baseline

To benchmark our model, the dataset is also utilized in replicating the methodologies outlined in the references FreeSense, Wi-alarm, and HAR.

FreeSense¹⁷: FreeSense identifies human motion by detecting the phase difference of amplitude waveforms on multiple antennas.

Wi-alarm¹⁸: Wi-alarm uses raw channel state information for human motion monitoring and uses SVM for detection.

HAR²¹: In this study, the author used CNN to perform edge detection on CSI data, enhancing human activity recognition based on WiFi.

Performance evaluation

Overall performance evaluation: To effectively assess its overall performance, the method was compared with FreeSense¹⁷, Wi-alarm¹⁸, and HAR²¹. As shown in Figs. 3 and 4, the method achieved the lowest FPR value across different data types, averaging approximately 1.2%, with a high TPR of 99.5%. This indicates fewer false alarms triggered by the method when detecting human presence.

Performance analysis in different environments: The versatility of our approach was further substantiated through a series of experiments conducted in diverse environments and postures, as depicted in Fig. 5. These figures illustrate the influence of environmental conditions on performance metrics. In both conference room and laboratory settings, the average TPR for detecting human presence was recorded at 98.8% and 98.4%, respectively, while the average FPR was 1.3% and 1.5%. This indicates a minor variance in the accuracy of human perception technology across different testing environments. The more spacious conference room experienced a reduced multipath effect. Conversely, the laboratory, smaller in size and cluttered with numerous objects, exhibited



Figure 3. Overall performance evaluation FPR.



Figure 4. Overall performance evaluation TPR.



Figure 5. FPR under different environment and personnel posture.

stronger multipath interference, leading to a decline in signal quality. Nevertheless, the overall TPR consistently exceeded 96%. Notably, the SNR in the conference room was approximately 19.7 dB, compared to 17.4 dB in the laboratory environment, underscoring the method's resilience under varying conditions.

Performance analysis across different body types: It allowed for the perception of individuals with varying genders, weights, and heights. The experimental results showed no substantial changes due to the posture of the subjects, thereby emphasizing the strong versatility of the method. As shown in Fig. 5.

Performance at different movement speeds: This study also assessed human motion perception at different speeds, such as slow walking, normal walking, fast walking, and sprinting. The experimental results were compared side by side with two typical methods, as shown in Figs. 6 and 7. The findings indicated that the method maintained a stable TPR and FPR across different movement speeds. However, the FPR significantly increased



Figure 6. TPR at different moving speeds.



Figure 7. FPR at different moving speeds.

while the TPR notably decreased for the other three methods when participants moved at slower speeds. This outcome can be attributed to the reduced interference characteristics of the wireless signal during the slower speed sampling period, causing less impact of the human body on the wireless signals. Nevertheless, the Self-Organizing Competitive Neural Network used in this method, with its rich feature set, maintained stable TPR and FPR, proving the robust reliability of the method.

Performance under environmental changes: To investigate the robustness of our model in the face of environmental changes, especially the effect of furniture rearrangement in rooms, we conducted targeted experiments. We modified the placement of chairs and tables in both lab and conference room settings to evaluate the impact on model performance. The results revealed that furniture rearrangement indeed influenced the model's performance. In the lab setting, post-adjustment, the TPR decreased by about 1.8%, and in the conference room, it decreased by approximately 2.1%. These findings suggest that while the model demonstrates considerable robustness, it is still somewhat affected by changes in furniture layout. Importantly, these tests were conducted without retraining the model, underscoring its adaptability to environmental shifts. However, the impact of furniture layout on performance warrants attention. In future research, we aim to explore this issue more thoroughly, with the goal of developing a more stable and efficient model for human presence detection. Through optimization and adjustments, we aspire to improve its detection accuracy and robustness in diverse settings, enhancing its applicability in smart homes, healthcare monitoring, and other scenarios.

Conclusion

Addressing the stability issues and user inconveniences of traditional human perception recognition techniques, this study presents a non-contact human presence detection technology. By preprocessing, extracting features, and classifying the CSI signals, we can discern different states such as an empty room, a room with a present individual, and a room where someone has recently been, demonstrating robust accuracy and versatility. Nevertheless, the current study only detects human presence within a room and does not recognize specific movements or simultaneous actions of multiple people, which limits its applicability. Future research will pivot towards the perception of multiple individuals' actions, broadening the scope and functionality of this technology.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Code availability

The code can be obtained from the corresponding author upon a reasonable request.

Received: 23 July 2023; Accepted: 8 February 2024 Published online: 13 February 2024

References

- 1. Zhao, H. *et al.* Metasurface-assisted massive backscatter wireless communication with commodity Wi-Fi signals. *Nat. Commun.* **11**(1), 3926 (2020).
- Zhu, X., Huang, Y., Wang, X. & Wang, R. Emotion recognition based on brain-like multimodal hierarchical perception. *Multimedia Tools Appl.* 2023, 1–19 (2023).
- Del Hougne, P., Fink, M. & Lerosey, G. Optimally diverse communication channels in disordered environments with tuned randomness. *Nat. Electron.* 2(1), 36–41 (2019).
- 4. Giunchiglia, F., Erculiani, L. & Passerini, A. Towards visual semantics. SN Comput. Sci. 2(6), 1–17 (2021).
- Wang, H., Cen, Y. & Zeng, X. Highly sensitive flexible tactile sensor mimicking the microstructure perception behavior of human skin. ACS Appl. Mater. Interfaces 13(24), 28538–28545 (2021).
- Bhalla, K. *et al.* Fusion of infrared and visible images using fuzzy based siamese convolutional network. *Comput. Mater. Contin.* 70(3), 5503–5518 (2022).
- 7. Halperin, D. *et al.* Tool release: Gathering 802.11n traces with channel state information. *ACM Sigcomm. Comput. Commun. Rev.* **41**(1), 53–53 (2011).
- 8. Del Hougne, P. et al. Dynamic metasurface aperture as smart around-the-corner motion detector. Sci. Rep. 8(1), 6536 (2018).
- 9. Dupre, M., del Hougne, P., Fink, M., Lemoult, F. & Lerosey, G. Wave-field shaping in cavities: Waves trapped in a box with controllable boundaries. *Phys. Rev. Lett.* 115, 017701 (2015).
- 10. Jin, Z. et al. WiFi-ID: Human Identification Using WiFi Signal. In International Conference on Distributed Computing in Sensor Systems (IEEE, 2016).
- 11. Cong, S. et al. Smart user authentication through actuation of daily activities leveraging WiFi-enabled IoT. In *The 18th ACM International Symposium* (ACM, 2017).
- 12. Faqiri, R. *et al.* PhysFad: Physics-based end-to-end channel modeling of RIS-parametrized environments with adjustable fading. *IEEE Trans. Wirel. Commun.* **22**(1), 580–595 (2022).
- 13. Del Hougne, P. Robust position sensing with wave fingerprints in dynamic complex propagation environments. *Phys. Rev. Res.* 2(4), 043224 (2020).
- 14. Hoomod, H. K., Al-Mejibli, I. & Jabboory, A. I. Analyzing study of path loss propagation models in wireless communications at 0.8 GHz. J. Phys. Conf. Ser. 1003(1), 458 (2018).
- 15. Suzuki, T. *et al.* Mechanical low-pass filtering of cells for detection of circulating tumor cells in whole blood. *Anal. Chem.* **92**(3), 2483–2491 (2020).
- 16. Li, H. Y. et al. Intelligent electromagnetic sensing with learnable data acquisition and processing. Patterns 1(1), 859 (2020).
- 17. Xin, T. *et al.* FreeSense: A robust approach for indoor human detection using Wi-Fi signals. *Proc. ACM Interact. Mobile Wear. Ubiquitous Technol.* **2**(3), 1–23 (2018).
- 18. Wang, T. et al. Wi-Alarm: Low-cost passive intrusion detection using WiFi. Sensors 19(10), 2335 (2019).
- 19. Liang, H. Evaluation of fitness state of sports training based on self-organizing neural network. *Neural Comput. Appl.* **33**(9), 3953–3965 (2021).
- 20. Lu, J. et al. SOFT: Softmax-free transformer with linear complexity. Adv. Neural Inf. Process. Syst. 34, 21297-21309 (2021).

- 21. Shahverdi, H. et al. Enhancing CSI-based human activity recognition by edge detection techniques. Information 14(7), 404 (2023).
- Del Hougne, P. & Lerosey, G. Leveraging chaos for wave-based analog computation: Demonstration with indoor wireless communication signals. *Phys. Rev. X* 8(4), 041037 (2018).
- Halperin, D., Hu, W., Sheth, A. & Wetherall, D. Predictable 802.11 packet delivery from wireless channel measurements. ACM SIGCOMM Comput. Commun. Rev. 40(4), 159–170 (2010).
- 24. Adib, F. & Katabi, D. See through walls with WiFi!. In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM* 75–86 (2013).
- Li, T., Fan, L., Yuan, Y. & Katabi, D. Unsupervised learning for human sensing using radio signals. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision 3288–3297 (2022).
- Fan, L. et al. Learning longterm representations for person re-identification using radio signals. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 10699–10709 (2020).
- 27. Zuo, J. et al. A new method of posture recognition based on wifi signal. IEEE Commun. Lett. 25(8), 2564–2568 (2021).
- Zhu, X. et al. An environmental intrusion detection technology based on WiFi. Wirel. Pers. Commun. 119(2), 1425–1436 (2021).
 Zhao, M., Adib, F. & Katabi, D. Emotion recognition using wireless signals. In Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking 95–108 (2016).
- Adib, F., Kabelac, Z. & Katabi, D. Multi-person localization via RF body reflections. In 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15) 279–292 (2015).

Author contributions

In this study, Y.Z. devised the main concept and drafted the paper, while also handling reviewer comments. J.W. was integral in data collection, processing, and the implementation of a neural network algorithm. X.W. focused on the utilization of machine learning techniques for state classification and recognition, contributing to the methodological writing and results interpretation. Lastly, X.Z. supervised the project, offering valuable insights into the study design and manuscript, and ensured the integrity and accuracy of the final paper.

Funding

This research work was not funded by any organization/institute/agency.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2024