# A Review of WiFi Sensing-based Gait Recognition from A Deep Learning Perspective

Yifei Chen *

College of Computer and Information Engineering, Nanjing Tech University, Nanjing, China

* Corresponding author: 202121139050@njtech.edu.cn

**Abstract.** Gait recognition based on WiFi sensing has emerged as a prominent area of research, driven by its potential for a diverse range of applications, the ability to recognize identity without being intrusive, and its cost-effectiveness. Deep learning models have been widely used in this recognition method, and different researchers have proposed the use of different neural networks to address different aspects of model-based gait recognition. However, the extant literature on these issues is lacking in a systematic compilation and evaluation. An overview of recent developments in deep learning models for WiFi-sensing-based gait identification is provided in this paper. It highlights key issues such as the removal of environmental dependency, the ability to deal with environmental random noise, and the handling of complex CSI data. Furthermore, research is conducted into the utilisation of multimodal recognition methods. In conclusion, the paper presents prospective solutions to the aforementioned issues and proposes avenues for future research.

**Keywords:** Deep learning, gait recognition, device-based sensing, machine learning.

## 1. Introduction

Gait is one of the most distinctive characteristics of an individual, and is generally determined by an array of factors including height, weight, limb length, and walking habits. These factors contribute to the unique and unchanging manner in which an individual walks. Gait recognition has emerged as a prominent research area, with applications spanning identity information recognition, disease diagnosis, and beyond. As an identity information recognition technology, it offers several advantages, including high security, low susceptibility to deception, non-invasive recognition, and a vast array of potential applications.

WiFi sensing is an emerging wireless sensing technology. This technology enables numerous human activity recognition exploits such as posture recognition, gesture recognition, and gait recognition by using WiFi signals to sense the surrounding environment. Gait recognition through WiFi sensing has the advantages of being independent of light conditions and targeting sensitive information with high privacy compared to computer vision solutions. It can utilize existing WiFi devices with few hardware updates, which reduces the deployment complexity and cost and improves the reachable range. It is more conducive to accurate recognition due to its bandwidth and signal processing capabilities [1].

Ahmad's review briefly explains the framework of human activity-based WiFi sensing [2]. WiFi-based sensing technologies work by capturing wireless channel disruptions caused by human activity and nearby movement, which in turn affects the propagation characteristics of the signal. Channel state information (CSI) modifications are what define this process [3]. CSI data can provide specific information about human movement and interaction behaviour. WiFi signals from WiFi routers, Intel 5300 Network Interface Cards (NIC), [3]and Atheros devices are particularly noteworthy due to their cost-effectiveness and low deployment difficulty [4,5]. Through the use of Orthogonal Frequency Division Multiplexing (OFDM) technology, these devices are incorporated into pre-existing WiFi infrastructures and divide the signal into several subcarriers, each of which sends data separately [6]. Human activity can cause interference to these subcarriers, which can change the characteristics of the signal captured by the CSI.

WiFi perception technology makes use of an extraction approach that extracts patterns from OFDM subcarriers to help identify and categorize human activities. It is a widely utilised technology in the field of context awareness. In contrast to conventional methodologies that depend on received signal strength indicators (RSSI), CSI techniques are capable of capturing the amplitude and phase nuances of signals, enabling the detection of subtle movements on channels of varying frequencies. Conventional methodologies are less effective in the presence of noise and minor movements. The technique comprises three phases: data collection, signal processing and classification. This structure allows the technique to offer the advantages of both accuracy and complexity.

WiFi sensing methods that are now in use include both deep learning-based and model-based methods. Model-based approaches do not produce precise identification results, but they are more resilient to changes in the environment. In the meanwhile, creating a mathematical model that links intricate human behavior to CSI shifts is exceedingly challenging. Occlusions such as walls and other obstacles can hinder the transmission of signals, making model-based techniques less effective in non-line-of-sight settings [7]. In these scenarios, deep learning-based methods have significant advantages. Based on this, a variety of WiFi sensing-based gait recognition models using deep learning methods are discussed in this paper.

## 2. Initial Processing of Deep Learning in WiFi Sensing-Based Gait Recognition

The signal processing process is the conversion of raw signal data into a description of human activity. It involves denoising the CSI data to obtain clearer CSI data, revealing temporal and frequency patterns through signal transformations (typically Fourier and Wavelet Transforms), and finally extracting these patterns by means of features for behavior recognition.

The preprocessing of CSI data represents a crucial stage in the context of WiFi-based human sensing. First, systematic errors are corrected by data calibration to ensure the accuracy of the measurements. Subsequently, denoising is performed using techniques such as filtering and smoothing to minimize the effects of environmental noise and fluctuations [8]. Outliers and abnormal data are also identified and removed at this stage to improve the quality of the data and lay the foundation for subsequent processing.

After preprocessing, signal transformation is a critical step in revealing underlying patterns and dynamic properties in CSI data. When converting data from the time domain to the frequency domain, the Fourier Transform (FT) is frequently utilized to resolve the frequency components connected to human behavior. In contrast, the wavelet transform helps to capture fast-changing transient events by simultaneously localizing signal features in time and frequency through multiresolution analysis [9].

Feature extraction obtains key information related to human activity from transformed data. Commonly used techniques include statistical feature extraction (e.g., mean, standard deviation, etc.), temporal feature analysis, and frequency domain feature extraction. To reduce noise and redundancy, Principal Component Analysis (PCA) and Independent Component Analysis (ICA) are commonly used for data dimensionality reduction, and these features provide the basis for classification in machine learning [10].

After completing signal processing, deep learning models are used to train and predict associations between specific signal patterns and human activities. These models were trained on big data and tested on unseen data to evaluate accuracy and robustness. The trained models were then deployed to accurately predict human activity based on WiFi signals.

## 3. Method and Application of Gait Recognition Based on WiFi Sensing Using Deep Learning

Ali presented a novel CNN-based technique that uses received wireless signal channel state information, filters outliers with the Hampel filter, eliminates noise and extracts significant features with the Discrete Wavelet Transform, and then uses machine learning and deep learning algorithms to identify the walking direction of various individuals in various environments [11]. When recognizing human movement directions in untrained data gathered in a classroom, a conference room, and two rooms, respectively, this technique obtained 92.9%, 95.1%, and 89% accuracy. The article demonstrates that deep learning can help improve the stability of gait recognition in different environments.

While randomly obtained CSI data are easily gathered, it is not practicable to gather high-quality CSI samples for adaptive algorithms in novel contexts due to the extreme environment reliance on many deep learning models. An efficient and annotated Wi-Fi sensing model based on a geometric self-supervised learning method was developed by Yang et al. named AutoFi [12]. Using randomly selected unlabeled low-quality CSI samples, AutoFi fully utilizes the signals before assigning them to user-specified tasks. For Wi-Fi sensing, this is the first study to enable cross-task transfer. By utilizing data from CSI samples that are randomly gathered, AutoFi achieves state-of-the-art performance in human gait identification. The self-supervised (GSS) learning module and the small-sample calibration (FSC) module make up AutoFi's two modules. Two random views are produced by augmenting randomly obtained CSI data in the self-supervised learning module. These views are then put into a feature extractor and a nonlinear function, which yields two distributions. The GSS loss forces the two predictive distributions to be consistent, which does not require any annotation. The trained feature extractor can then be transferred to the FSC module. For identity recognition, a 27.25% improvement in accuracy was achieved in the complex dressing scenario in comparison to a single prototype network. AutoFi significantly improves small sample performance, enhances the existing system through cross-task knowledge transfer, and makes deep learning models a significant step forward in WiFi sensing gait recognition in scenarios where no developers are involved.

Random noise often occurs indoors in real recognition scenarios. A large number of current identification techniques have poor identification accuracy and are vulnerable to random noise in indoor situations. A device-less CSI human identification approach based on deep learning called Wihi was suggested by Ding [13]. To identify various persons, Wihi primarily uses three important strategies. The discrete wavelet transform technique is first used to denoise the original CSI data using signal decomposition to remove the impact of random noise. Second, several representative features are taken from various statistical profiles, such as energy distribution across multiple frequency bands, time-frequency analysis, and channel power distribution in the time domain, to thoroughly define human gait. Ultimately, the aforementioned typical gait parameters are learned by a recurrent neural network (RNN) model equipped with long and short-term memory blocks, which then encode temporal information necessary for human identification. According to experimental data, Wihi performs better in the laboratory when random noise is present than current approaches, which makes it perfect for indoor large-scale deployment.

Most deep learning models based on WiFi perception are still stuck in traditional machine learning methods due to limitations such as data structure. With the development of hardware technology, there are more and more subcarrier sequences that can be captured in CSI, making the CSI matrix more like an image than a time series. Ding et al. believe that the accuracy of the model can be further improved by drawing on deep learning methods in the fields of image processing and computer vision [14]. The team collected more than 1,000 pieces of data from three volunteers, added the attention mechanism to the residual neural network pre-trained with ImageNet, and obtained a better model, Vi-WiFi-Gate, after several trainings. In multi-target classification, the deep neural network of the article achieved 94.6% recognition accuracy, which proves the feasibility of borrowing the cross-domain deep learning methods.

In low light, vision-based gait identification techniques perform less well because they are unreliable. GaitFi is a brand-new multimodal gait identification technique that Deng et al. suggested [15]. Videos and WiFi signals are used in this strategy to facilitate human identification. To capture human gait, GaitFi gathers channel state data representing WiFi multipath propagation. A video camera then records the video. Deng suggests a Lightweight Residual Convolutional Network (LRCN) as the backbone network and furthermore suggests a dual-stream GaitFi in order to get robust gait information. Achieving the gait retrieval problem involves combining visual and WiFi characteristics. GaitFi is trained on different levels of features through triple loss and classification loss. Extensive real-world experiments were conducted by Deng et al. The findings of the trial show that GaitFi works better than the most advanced gait identification techniques relying on a single WiFi or camera. GaitFi achieves an accuracy of 94.2% in a human body recognition task with 12 subjects. The proposed model establishes the foundation for deep learning models to be applied to multimodal gait recognition methods.

## 4. Challenges and Prospects

In the context of WiFi-based sensing systems, the impact of changes in the physical environment represents a significant challenge. To illustrate, the rearrangement of furniture or the introduction of obstacles may result in inconsistencies between the received data and the original reference data, which in turn affects the perception. Even though these environmental changes are common, they might cause the sensor system's performance to deteriorate. Consequently, in order to guarantee the system's resilience and dependability in dynamic real-world settings, it is imperative to design adaptive algorithms and deep learning approaches that can quickly adjust to such environmental changes.

Differences in technical standards, operating principles, and frequencies for multi-user sensing in multi-modal scenarios bring more complexity to the research. Currently, many studies focus on single RF modes, but WiFi signals can be used in combination with other RF modes, such as FMCW radar (operating in millimeter-wave and terahertz bands) and LoRa signals. In this multi-modal environment, how to effectively train deep learning models to support multi-user perception becomes a direction worth exploring in depth.

Edge intelligence applications of deep learning algorithms for WiFi sensing still face many challenges. First, the computational requirements of deep learning models make it difficult to perform real-time inference on resource-constrained embedded devices, limiting their usefulness in various scenarios. Second, these models usually require more memory and storage resources, which puts greater demands on embedded devices. In addition, the transparency and interpretability of deep learning models also remain deficient, and it is crucial to ensure their transparency when dealing with sensitive WiFi data. Therefore, exploring memory-efficient techniques and optimization strategies is crucial to promote the wide application of deep learning-based WiFi sensing systems and enhance the performance of human activity recognition.

## 5. Conclusion

This paper reviews the research results of gait recognition based on WiFi sensing and discusses the application of deep learning models in this field. The findings of the trial show that GaitFi works better than the most advanced gait identification techniques relying on a single WiFi or camera. The combination of traditional visual recognition and WiFi sensing recognition technology to achieve bimodal recognition can complement each other's advantages, meet the identification of identity information under different environmental changes, and greatly improve recognition accuracy. Convolutional neural networks (CNN) and long short-term memory networks (LSTM) are two popular deep learning models that are utilized for extracting gait features and navigating challenging settings. However, the robustness and adaptability of existing techniques in complex scenarios still need to be improved. The accuracy and real-time performance of WiFi gait recognition will be further enhanced in the future with the development of hardware and multimodal data fusion technology, particularly in the highly promising fields of smart home, intelligent security, and health monitoring.

## References

[1] Liu, Z., Liu, X., Zhang, J., Li, K. Opportunities and challenges of wireless human sensing for the smart IoT world: A survey. IEEE Network, 2019, 33 (5), 104 - 110.

[2] Ahmad, I., Ullah, A., Choi, W. WiFi-Based Human Sensing with Deep Learning: Recent Advances, Challenges, and Opportunities. IEEE Open Journal of the Communications Society. 2024.

[3] Wang, Z., Huang, Z., Zhang, C., Dou, W., Guo, Y., Chen, D. CSI-based human sensing using model-based approaches: a survey. Journal of Computational Design and Engineering, 2021, 8 (2): 510 - 523.

[4] Halperin, D., Hu, W., Sheth, A., Wetherall, D. Tool release: Gathering 802.11 n traces with channel state information. ACM SIGCOMM computer communication review, 2011, 41 (1): 53 - 53.

[5] Xie, Y., Li, Z., Li, M. Precise power delay profiling with commodity WiFi. In Proceedings of the 21st Annual International Conference on Mobile Computing and Networking 2015.

[6] Li, H., Yang, W., Wang, J., Xu, Y., Huang, L. WiFinger: Talk to your smart devices with finger-grained gestures. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing,2016.

[7] Wang, Z., Huang, Z., Zhang, C., Dou, W., Guo, Y., & Chen, D. CSI-based human sensing using model-based approaches: a survey. Journal of Computational Design and Engineering, 2021, 8 (2): 510 - 523.

[8] Santhalingam, P. S., Hosain, A. A., Zhang, D., Pathak, P., Rangwala, H., & Kushalnagar, R mmASL: Environment-independent ASL gesture recognition using 60 GHz millimeter-wave signals. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4 (1): 1 - 30.

[9] Ge, Y., Taha, A., Shah, S. A., Dashtipour, K., Zhu, S., Cooper, J., Imran, M. A. Contactless WiFi sensing and monitoring for future healthcare-emerging trends, challenges, and opportunities. IEEE Reviews in Biomedical Engineering, 2022, 16, 171 - 191.

[10] Patil, M. N., Iyer, B., & Arya, R. Performance evaluation of PCA and ICA algorithm for facial expression recognition application. In Proceedings of Fifth International Conference on Soft Computing for Problem Solving: SocProS, 2016, 1: 965 – 976.

[11] Ali, H. A. H., Seytnazarov, S. Human Walking Direction Detection Using Wireless Signals, Machine and Deep Learning Algorithms. Sensors, 2023, 23 (24): 9726.

[12] Yang, J., Chen, X., Zou, H., Wang, D., Xie, L. Autofi: Toward automatic wi-fi human sensing via geometric self-supervised learning. IEEE Internet of Things Journal, 2022, 10 (8): 7416 - 7425.

[13] Ding, J., Wang, Y., Fu, X. Wihi: WiFi-based human identity identification using deep learning. IEEE Access, 2020,8: 129246 - 129262.

[14] Li, T., Chen, T., Li, K., Ao, R. Vi-WiFi-Gate: a WiFi sensing gait recognition method inspired by the image vision field. In International Conference on Internet of Things and Machine Learning, 2023, 12937: 376 – 381.

[15] Deng, L., Yang, J., Yuan, S., Zou, H., Lu, C. X., Xie, L. GaitFi: Robust device-free human identification via WiFi and vision multimodal learning. IEEE Internet of Things Journal, 2022 10 (1): 625 - 636.