Analysis of Human Behavior Based on Deep Learning Models

Hongye Hu

School of Information Engineering, Guangdong University of Technology, Guangzhou,510000, China

3122002223@mail2.gdut.edu.cn

**Abstract.** With the development of Internet of Things (IoT) technology and the widespread popularity of smart devices, wireless sensing has received widespread attention as an emerging means of human behaviour recognition. By using the existing Wi-Fi signals in the environment instead of arranging additional sensors, it not only reduces the cost but also improves the covertness and user acceptance of the monitoring system. In this study, an efficient and real-time human behaviour recognition model is constructed based on wireless sensing technology, using an Intel 5300 network card to capture channel state information (CSI) and combining LSTM and CNN deep learning models combined for data processing. The system performs well in the task of classifying human behaviours, especially demonstrating the recognition accuracy of different human behaviours. After the number of training times reaches 50, the average recognition accuracy reaches 95%, and the value of the loss function effectively decreases with the increase of the number of training times, and the model fits well to the dataset, proving the effectiveness and stability of the model.

# Introduction

With the rapid development of smart home, health monitoring, security system and human-computer interaction, behavior recognition technology has become one of the key technologies to realize an intelligent society[1,2]. The rise of wireless sensing technology brings new opportunities and challenges to this field. Compared with the traditional behavior recognition methods based on visual or wearable sensors, wireless sensing technology not only reduces the cost but also improves the covertness and user acceptance of the monitoring system by using the existing Wi-Fi signals in the environment instead of arranging additional sensors. This approach is particularly suitable for scenarios with high privacy requirements and provides reliable performance in complex environments. Therefore, how to use wireless sensing technology, especially by analyzing the channel state information (CSI) in Wi-Fi signals to accurately detect the environment and human dynamics for effective recognition of human behavioral patterns, has become an issue worthy of in-depth exploration[3].

In recent years, many scholars have devoted themselves to exploring the application of deep learning models in the context of wireless sensing, especially the application of Long Short-Term Memory Networks (LSTMs) and Convolutional Neural Networks (CNNs), which are widely regarded as powerful tools for processing time series data and spatial feature extraction. For example, some scholars exhibited how to use LSTM models combined with CSI data for indoor localization and tracking, and their results showed that the method can significantly improve the localization accuracy[4-6]. Other scholars propose a CNN-based framework for analyzing CSI data, which is used to distinguish different human activities and achieve high-accuracy activity recognition. In addition, other scholars detailed how to use Intel 5300 NIC to capture CSI data and apply it to the fall detection system for the elderly, which proved the potential of this technology in health monitoring. Not only that, but there are also scholars who can further improve the accuracy and stability of behavior recognition by optimizing the CSI data capture method and combining it with advanced machine learning algorithms. These studies show that combining deep learning with CSI data can effectively improve the performance and usefulness of behavior recognition systems.

The purpose of this paper is to explore how to construct an efficient and real-time human behavior recognition system by integrating LSTM and CNN deep learning models and utilizing CSI data captured by Intel 5300 NICs. In this paper, we first review the related technical background and existing research results, and then detail our proposed system architecture and its working principle. Next, we show the experimental design, dataset description, evaluation metrics, and analysis of experimental results. Finally, this paper discusses the limitations of the system and provides an outlook on possible future research directions. Through this research, we hope to provide new perspectives and solutions for the application of wireless sensing technology in the field of behavior recognition, and at the same time promote its implementation and development in several practical application scenarios, such as smart home and health monitoring.

# Data and methods

## Data

This dataset contains raw CSI data from a range of behavioral activities performed by 8 different volunteers (v1 to v8). These data are saved in DAT file format, with each file name encoding specific information for easy identification and processing. v: represents the volunteer number, ranging from 1 to 8 (i.e., v1, v2,... , v8), indicating a total of 8 different participants. a: represents the behavior number, ranging from a1 to a10, indicating that each volunteer performed a total of 10 different types of behaviors[7-8].

The last digit represents the number of times the same behavior was performed repeatedly, which helps to increase the amount of data and provide multiple measurements to ensure the reliability and accuracy of the data. All data was captured via an Intel 5300 network card, a device capable of capturing detailed channel state information (CSI).CSI data reflects how a wireless signal travels from the transmitter to the receiver, including information on path loss, shadowing effects, and multipath propagation, which is critical to understanding environmental characteristics and human dynamics.

Because there were eight volunteers, each of whom performed 10 different behaviors and at least multiple repetitions of the experiment for each behavior, the entire dataset contains a large number of rich CSI data points, providing a solid foundation for in-depth analysis.

In order to improve the quality of the data and ensure its consistency, accuracy, and usability, this paper preprocesses the data so as to enhance the effect of subsequent analysis or model training. The first step is loading and filtering. Time series data in .npy format are loaded from a specified directory, and these data are filtered using an 8th order low-pass filter. The cutoff frequency of the filter is set to 0.5 (normalized frequency) to remove high frequency noise[9]. Next, the data is segmented by dividing the filtered data into 5-second segments. For each 5-second segment, the middle 2 seconds are extracted as the valid part of the data. Finally, the data is transformed, which includes the dimensionality reduction by PCA (Principal Component Analysis), and the design plan is to reduce the valid data segments of each 2 seconds to 20 principal components by PCA. The data were subjected to transposition operations to change the dimensional order of the data to accommodate possible subsequent analysis.

In the data preprocessing procedure of this study, the quality of the CSI data was optimized through a series of steps to ensure that it was more suitable for subsequent analysis and model training. First, a low-pass filter was applied to effectively remove high-frequency noise from the signal, which improved the accuracy of the subsequent analysis[10]. Next, by selecting the middle 2 seconds of each 5-second data segment as the effective portion, focusing on the most stable data segments, the effects of non-stationary or transitional states at the beginning and end phases were avoided. In addition, the use of the PCA dimensionality reduction technique not only simplifies the computational complexity, but also preserves the important feature information, which facilitates the subsequent deep learning model for more efficient processing. Finally, the entire preprocessing process is implemented through automated scripts for systematic data management. This approach ensures that every step from data capture to final analysis can be carried out more precisely and efficiently. Such a one-stop processing flow not only reduces the possibility of human error, but also speeds up the transformation from raw data to valuable insights.

## Methodology

In the experiments, an Intel 5300 NIC was chosen as the hardware platform in order to capture channel state information (CSI). The Intel 5300 NIC is able to provide detailed channel state information in wireless network environments, which is crucial for Wi-Fi-based sensing and recognition tasks. With this NIC, high-precision CSI data containing information on multipath effects, signal fading, etc., can be acquired, which provides high-quality input for subsequent data analysis and model training. For the selection of deep learning models, Long Short-Term Memory Network (LSTM) and Convolutional Neural Network (CNN) are used. Each of these two models has unique advantages and is suitable for processing complex sequential data.

An important feature of the Intel 5300 NIC, a wireless NIC that supports the 802.11n standard, is its ability to capture detailed CSI data.CSI data describes how a signal travels from the transmitter to the receiver, and includes information such as path loss, shadowing effects, and multipath propagation. This information is crucial for understanding the behavior of wireless channels and has important applications in WiFi based motion recognition, indoor localization, and other areas.

LSTM is a special type of recurrent neural network (RNN) specifically designed to solve long-term dependent problems. This means that LSTM is well suited for processing time-series data such as text, speech, or CSI data, as mentioned here. Through its unique gating mechanism, LSTM is able to selectively remember or forget information, which makes it more effective than traditional RNNs in processing information with long time spans.

CNNs are mainly used to process data with a grid-like structure, such as image data. However, CNNs are also effective at capturing localized features and patterns when processing sequence data. When applied to one-dimensional sequence data, CNNs can detect patterns in the sequence through a filter sliding window, which is useful for extracting specific features in CSI data.

Using LSTMs in combination with CNNs can provide advantages in processing complex sequence data.

The first is an enhanced feature extraction capability; CNNs are good at automatically extracting local features of the input data, which are often key to understanding the data. In CSI data analysis, CNNs can be used to identify specific patterns or anomalies in the data. Afterwards, LSTM can then use these features to further analyze their trends over time, thus improving the overall recognition accuracy.

Second is the ability to handle long sequence data, while CNNs perform well with short sequences or fixed-length inputs, they may encounter challenges when dealing with long sequence data.LSTM is able to compensate for this shortcoming as it is particularly suited to handle data with varying sequence lengths and is able to capture long-term dependencies between the data.

Finally, flexibility and adaptability are enhanced, and the combined use of CNN and LSTM allows the model to have both strong feature extraction capabilities and an in-depth understanding of time-series data. This combination not only improves the expressive power of the model, but also increases its adaptability to different types of data and application scenarios, especially when spatial and temporal information need to be considered together.

In summary, the hybrid model constructed by using an Intel 5300 NIC for CSI data capture and combining CNN and LSTM can effectively improve the performance of CSI data-based applications, such as action recognition or indoor localization tasks, in terms of accuracy.

## Indicators

Accuracy is the ratio of the number of correctly categorized samples to the total number of samples. It is a basic indicator of the overall performance of the model.

The confusion matrix visually presents the model's identification of various types of behaviors, with special emphasis on the categories of behaviors that are easily confused.

The loss function used during training is an important metric for evaluating the model's learning process. A lower loss value means that the model's predictions are closer to the true label.

The impact of environmental noise on the recognition accuracy is evaluated, and the robustness of the model in the real environment is verified.

# Analysis of results

The experiment used the CSI dataset captured by Intel 5300 NIC, and designed ten categories of tasks including a0-a9 standing; walking, bending; squatting, drinking; cleaning; running; sitting; lying down; and falling, respectively. The confusion matrix analysis demonstrates how well the model recognizes the various types of behaviors, especially pointing out the confusion point between 'fall' and 'lie down' behaviors.

At the early stage of training, i.e., when the number of iterations is 1, the accuracy (ACC) of the test set is only about 49.3%, while the accuracy of the training set is about 30.1%(Table 1). At this point, the model is still in the initial stages of learning from the data and the overall loss value is high. From iteration 1 to 25, the model gradually learns how to better recognize different human behaviors as the training progresses. By the 25th iteration, the accuracy of the test set significantly improved to about 93.61%, while the accuracy of the training set was even higher at about 98.36%. This indicates that after 25 rounds of training, the model has a good fit to the training dataset and is able to recognize almost all the training samples correctly. When the number of iterations reaches 50, the accuracy of the test set remains at a high level, although it drops slightly to about 89.9%. Meanwhile, the accuracy of the training set drops slightly to about 97.2%. Nonetheless, this change does not signify a decline in model performance, but instead may be due to the model starting to avoid overfitting the training data and thus performing slightly more conservatively in the face of unseen test data. After 25 training sessions, the model's recognition accuracy for human behavioral actions stabilizes at over 95%.

**TABLE 1:** Number of training sessions, Accuracy of training and test sets, and performance of training models with total loss values

|  |  |  |
| --- | --- | --- |
| Epoch | ACC | Total Loss |
| 1 | Test： 0.493 | 159.771 |
| Train： 0.301 |
| 25 | Test：0.936 | 114.767 |
| Train：0.983 |
| 50 | Test：0.899 | 114.966 |
| Train：0.972 |

# Conclusion

In this paper, we focused on the application of wireless sensing technology in human behavior recognition, and constructed an efficient and real-time human behavior recognition model by combining the channel state information captured by Intel 5300 NIC with deep learning models, LSTM and CNN.

Experimental results show that the system performs well in the behavior classification task. After several iterations of training, the recognition accuracy of the model is stable at more than 95%, especially after the number of training times reaches 25, the accuracy of the model on the test set reaches 93.61%, and the value of the loss function continues to decrease with the increase of the number of training times, which indicates that the model has a good generalization ability and stability. The main finding of this study is that the accuracy of behavior recognition based on CSI data can be significantly improved by integrating the advantages of both deep learning models, LSTM and CNN. Specifically, CNN is good at capturing the local features of the input data, while LSTM is able to effectively analyze the trend of these features over time, and the combination of the two not only enhances the feature extraction capability, but also improves the ability to handle long sequential data, which enables the model to perform well in complex environments.

Looking ahead, although our research has made significant progress, there is still room for improvement. For example, the performance of the system can be further enhanced by optimizing the data acquisition strategy or introducing more advanced machine learning algorithms. In addition, considering the diversity and dynamic changes in real-world application scenarios, how to enhance the adaptability of the model is also a direction worth exploring. The significance of this research not only lies in proposing a new method of behavior recognition, but also in providing technical solutions for a number of fields, such as smart home and health monitoring, which are expected to promote the development of technology and realize a more intelligent living environment. In conclusion, through continuous exploration, wireless sensing technology will show its potential in more realistic application scenarios.

# ReferenceS

1. F. Zhang, C. Wu, B. Wang, M. Wu, D. Bugos, H. Zhang, and K. J. R. Liu, SMARS: Sleep Monitoring via Ambient Radio Signals, IEEE Trans. Mobile Comput., 20(5), 1370–1380, (2021).
2. J. Liu, Z.-J. Zha, and R. Hong, Dual Context-Aware Refinement Network for Person Search, Proc. 28th ACM Int. Conf. Multimedia, 3450–3459, (2020).
3. H. El Zein et al., Leveraging Wi-Fi CSI Data for Fall Detection: A Deep Learning Approach, 5th Int. Conf. Bio-eng. Smart Technol. (BioSMART), 1–4, (2023).
4. G. Sheng and L. You, Indoor Positioning Method Based on Multi-User WiFi Signal Strength, Wireless Commun. Mobile Comput., (2020).
5. B. Yu, Y. Wang, K. Niu, Y. Zeng, T. Gu, L. Wang, C. Guan, and D. Zhang, WiFi-Sleep: Sleep Stage Monitoring Using Commodity Wi-Fi Devices, Proc. ACM Interact. Mobile Wearable Ubiquitous Technol., (2021).
6. K. Zhang, Q. Liang, and Q. Wu, HarFi: Human Trajectory Recognition Based on WiFi CSI Using Deep Learning, 9th Int. Conf. Comput. Commun., (2023).
7. X. Liu, J. Cao, S. Tang, and J. Wen, Wi-Sleep: Contactless Sleep Monitoring via WiFi Signals, IEEE Real-Time Syst. Symp., (2014).
8. P. Hu, W. Liu, C. Yang, Y. Sun, and J. Li, WiFi CSI Based Passive Human Activity Recognition Method Using BLSTM-CNN, 23rd Int. Conf. Commun. Technol., (2020).
9. H. Ji-ai, L. Xian-qi, and Z. Xue, A Human Activity Recognition Method in Indoor WiFi Environment, Commun., (2023).
10. Y. Wang, Y. Wang, Q. Liu, and Y. Zhang, Dynamic WiFi Indoor Positioning Based on the Multi-Scale Metric Learning, Comput. Commun., 213, 49–60, (2024).