J. Meas. Sci. Instrum., 2025, 16(4): 569-577. http://jmsi.publish.founderss.cn/
DOI: 10.62756/jmsi.1674-8042.2025055 jmsi@nuc.edu.cn

Indoor localization with channel state information images from selected multiple access points

LONG Liang, WANG Xiaopeng*, WANG Jiang, LI Gang

School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

***Corresponding author:** WANG Xiaopeng (Wangxiaopeng@mail.lzjtu.cn)

Received: August 11, 2023 **Revised:** October 20, 2023 **Accepted:** January 15, 2024

Abstract: To improve the accuracy of indoor localization methods with channel state information (CSI) images, a localization method that used CSI images from selected multiple access points (APs) was proposed. The method had an off-line phase and an on-line phase. In the off-line phase, three APs were selected from the four APs in the localization area based on the received signal strength indication (RSSI). Next, CSI data was collected from the three selected APs using a commercial Intel 5 300 network interface card. A single-channel sub-image was constructed for each selected AP by combining the amplitude information from different antennas and the phase difference information between neighboring antennas. These sub-images were then merged to form a three-channel RGB image, which was subsequently fed into the convolutional neural network (CNN) for training. The CNN model was saved upon completion of training. In the on-line phase, the CSI data from the target device was collected, converted into images using the same process as in the off-line phase, and fed into the well-trained CNN model. Finally, the real position of the target device was estimated using a weighted centroid algorithm based on the model's output probabilities. The proposed method was validated in indoor environments using two datasets, achieving good localization accuracy.

Key words: WiFi indoor localization; multiple access points; channel state information image; convolutional neural network (CNN); fingerprint localization; weighted centroid algorithm

0 Introduction

Indoor localization plays a significant role in various aspects of modern life, including indoor navigation, intelligent building management, and personalized services[1-6]. Among indoor localization technologies, WiFi-based methods have gained prominence due to their cost-effectiveness, ease of deployment, and scalability.

Many WiFi indoor localization methods currently use received signal strength (RSS) [7-10] for localization. However, in complex indoor environments, RSS is prone to the influence of factors such as multipath effects, shadow fading, and signal interference, thereby limiting the accuracy and reliability of localization. In contrast, channel state information (CSI) provides more granular physical layer information, such as the phase and amplitude of multi-channel subcarriers. This detailed information can more accurately reflect signal variations during transmission. Therefore, many methods use CSI for indoor localization[11-15]. However, CSI data does not directly provide helpful information about the wireless channel, so researchers have proposed using CSI images for indoor localization. By converting CSI data to images, researchers can comprehensively use the information in CSI while avoiding the subjective feature selection and preprocessing steps. Chen et al. [16] constructed CSI images using amplitude information from three antennas, achieving an average localization error of 1.37 m with a convolutional neural network (CNN) as the localization model. Zhu et al. [17] constructed CSI images of dimensions $30\times30\times3$ using phase information from three antennas, achieving an average localization error of 1.28 m with a CNN as the localization model. Wang et al. [18] used phase differences information between adjacent antennas to construct single-channel CSI images of size 60×60, achieving an average localization error of 1.78 m using a deep convolutional neural network as the localization model. Li et al. [19] combined amplitude and phase differences information of CSI to construct three-channel images of dimensions $114\times114\times3$, with one channel constructed by amplitude of an antenna and the other two channels constructed by phase differences between neighbouring antennas,

achieving an average localization error of 1.58 m. Liu et al.[20] achieved average localization errors of 0.28 m and 0.48 m in corridor and laboratory environments using amplitude and phase differences to construct CSI images of dimensions 30×30×3. However, the CSI image-based indoor localization methods mentioned above only use CSI from a single access point (AP). Consequently, they do not comprehensively consider CSI differences among different AP paths. Particularly in complex indoor environments, the impact of multipath effects can distort CSI data from one AP. However, CSI data from other APs can compensate for this distortion, significantly alleviating multipath effects. Furthermore, reducing the impact of APs with poor signal quality on localization is also significant.

We proposed a novel CSI image construction method that comprehensively described the signal characteristics in indoor environments by combining amplitude and phase differences information from selected multiple APs. By selecting three APs from four APs based on their received signal strength indication (RSSI) as the localization APs and collecting amplitude and phase differences data from all antennas, we constructed novel three-channel CSI images as the fingerprints for localization. This method comprehensively considered CSI differences along different paths and mitigated the impact of AP with the lowest RSSI on localization.

1 Relevant theory

1.1 Channel state information

Channel state information describes the signal propagation and reveals the effects of scattering and attenuation during signal transmission. It provides information about a communication link's channel characteristics, enabling the channel's estimation. In an orthogonal frequency division multiplexing (OFDM) system, the received signal after passing through a multipath channel can be expressed as

$$\boldsymbol{Y}=\boldsymbol{H}\boldsymbol{X}+\boldsymbol{N}, \tag{1}$$

where $\boldsymbol{Y}$ is the received vector, $\boldsymbol{X}$ is the transmitted vector, $\boldsymbol{H}$ is the channel matrix, $\boldsymbol{N}$ is the additive white Gaussian noise vector.

The channel matrix $\boldsymbol{H}$ records the characteristics of the wireless channel, including the amplitude and phase information of CSI. Commercial wireless network cards, such as the Intel 5300, integrate OFDM technology that supports the IEEE 802.11n standard, allowing the direct extraction of CSI from 30 subcarriers. The CSI of each subcarrier can be expressed as

$$\boldsymbol{H}_i=|\boldsymbol{H}_i|\exp\left(\mathrm{j}\angle\boldsymbol{H}_i\right), \tag{2}$$

where $|\boldsymbol{H}_i|$ is the amplitude of the subcarrier, and $\angle\boldsymbol{H}_i$ is the phase of the subcarrier.

1.2 Convolutional neural network

The core components of a convolutional neural network are the convolutional layer, the pooling layer, and the fully connected layer. The convolutional layer operates by applying convolution operations to the input image using a set of trainable filters known as convolutional kernels. These kernels are responsible for capturing local features within the image. Within the convolutional layer, multiple neurons process the output from the preceding layer before passing it through the activation function during forward propagation. Let $k_{i,j}$ denote the convolutional kernel as having a local weight matrix. Thus, the convolutional layer can be expressed as

$$Conv_{j,m}=A\left(\sum_{i=1}^{I}\sum_{n=1}^{P}o_{i,n+m-1}\,k_{i,j,n}+k_{o,j}\right), \tag{3}$$

where $o_{i,m}$ is the output of the i-th neuron of the m-th feature, $k_{i,j,n}$ is the n-th convolutional kernel, $k_{o,j}$ is the connection between the o-th and j-th feature maps, P is the size of the convolutional kernel, and $A(\cdot)$ is the activation function expressed as

$$A(x)=\begin{cases}x, & x<0,\\ 0, & x\geqslant 0.\end{cases} \tag{4}$$

The pooling layer and the convolutional layer have the same number of feature maps. However, the pooling layer can reduce the resolution of these feature maps, thereby facilitating the abstraction of lower-dimensional features. A commonly used pooling technique is max pooling, which can be expressed as

$$Pool_{i,m}=\max_{n=1}^{L}Conv_{i,(m-1)\times s+n}, \tag{5}$$

where L is the pooling size and s is the stride size.

The fully connected layer is commonly located in the last few layers of a CNN model and is used to transform the extracted features into the final output results. During the training process, the weights of the fully connected layer are updated using the backpropagation algorithm to minimize the loss function, enabling the neural network to learn appropriate feature representations and classification decisions for the task at hand. Cross entropy is commonly chosen as the loss function to measure the difference between the predicted results and the true labels. It is expressed as

$$E=-\sum_{j=1}^{T} y_i \log\ p_j, \tag{6}$$

where $\boldsymbol{y}$ is a vector consisting of elements 0 and 1, and p is the output of the softmax function.

1.3 Principle of fingerprint-based indoor localization

The fingerprint-based localization method aims to obtain the feature information of the measured location (the so-called fingerprint). In the real-time localization phase, the WiFi signal of the device to be localized is compared with the known features to estimate the position. Indoor localization based on CNN using CSI as a fingerprint involves the following phases. Firstly, it is necessary to collect CSI data within the indoor environment. This data contains signal amplitude and phase information and is divided into training and testing datasets. Data cleansing, noise filtering, and normalization can be applied during the phase. Subsequently, a CNN model is created to learn the mapping relationship from CSI data to location coordinates. The CNN model is trained using CSI data from the training dataset and corresponding known location information. The training objective is to minimize the error between predicted and actual positions.

Through the backpropagation algorithm, the model gradually adjusts weights and biases to adapt to the data's features. Following the completion of training, the CNN model retains a mapping relationship from CSI data to location coordinates. This model can be considered a unique CSI fingerprint, as it maps CSI data to indoor positions. CSI data from the current location is collected when device localization is needed. This data is fed into the well-trained CNN model, which predicts the device's position based on the learned mapping relationship.

2 Localization system

2.1 System structure

The proposed localization system has on-line and off-line phases. The off-line phase primarily involves constructing the CSI image and training the localization model, while the on-line phase primarily involves predicting the position of the target device. The proposed localization system is shown in Fig.1.

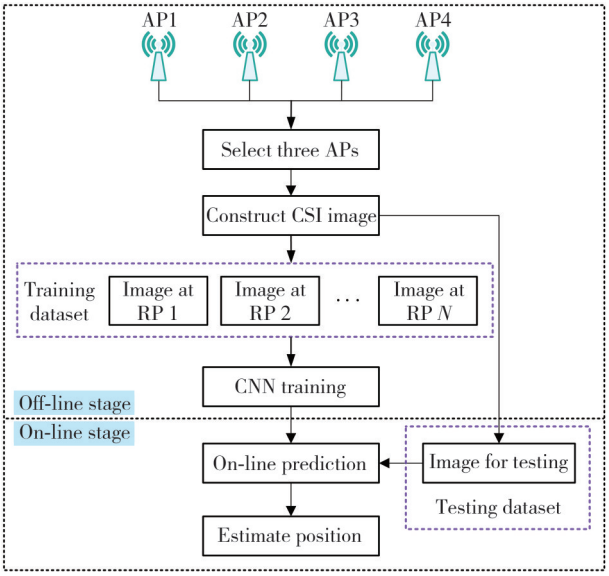


Fig. 1 System structure

2.2 Off-line phase

In the off-line phase, we deploy four single-antenna APs in the target localization area and set reference points (RPs) at even intervals of 0.5 m. For each RP, we measure the RSSI from each AP at its position and then select three APs with the highest RSSI from the four APs. Subsequently, we collect the CSI data from the three APs using a receiving device with three antennas and normalize its values to a range between 0 and 255. Finally, a novel three-channel CSI image is constructed by combining the CSI data's amplitude and phase differences information from the different antennas of the three selected APs. Each channel of the image is essentially a real-valued matrix with sizes of 180×180, and its composition can be defined as

$$Img_i=[\boldsymbol{A}_1;\boldsymbol{A}_2;\boldsymbol{A}_3;\boldsymbol{Pd}_1;\boldsymbol{Pd}_2;\boldsymbol{Pd}_3], \tag{7}$$

where Img_i is the image of the i-th AP, $i\in\{1,2,3\}$. $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_3$, $\boldsymbol{Pd}_1$, $\boldsymbol{Pd}_2$, and $\boldsymbol{Pd}_3$ are all matrices of size 180×30, where 180 rows represent 180 data packets, and 30 columns represent 30 subcarriers in each data packet. $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, and $\boldsymbol{A}_3$ are the amplitude information of CSI received by the first, second, and third antennas of the receiving device, respectively. $\boldsymbol{Pd}_1$ is the phase differences information of CSI between the first and second antennas of the receiving device. $\boldsymbol{Pd}_2$ is the phase differences information of CSI between the second and third antennas of the receiving device. $\boldsymbol{Pd}_3$ is the phase differences information of CSI between the third and first antennas of the receiving device. Fig.2 shows an example of a CSI image, where each sub-image (Img_1, Img_2, Img_3) is an 8-bit grayscale image. These sub-images ultimately merge to construct a three-channel RGB image.

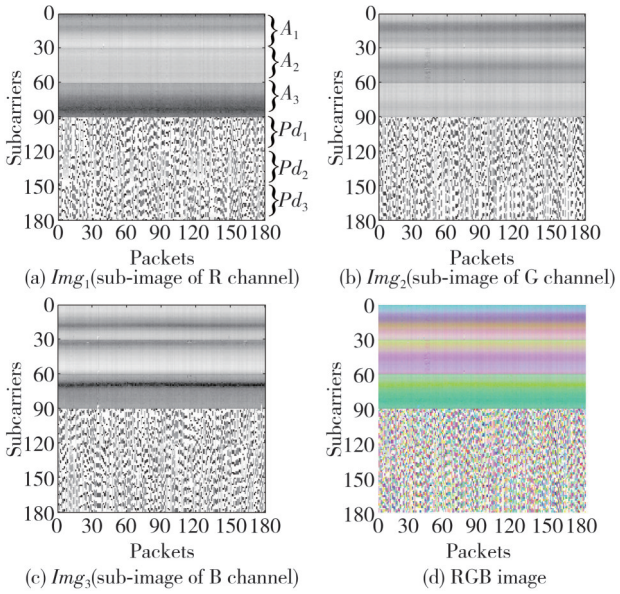


Fig. 2 Example of CSI image

As shown in Fig.2, although the CSI data is collected from the same RP position, the images from different APs show significant variations.

Furthermore, even for the images from the same AP, there is a clear difference between the local image of magnitude and the local image of phase differences.

Furthermore, the local images of magnitude or phase differences from different antennas are also different. These observations suggest that combining amplitude and phase differences information from different antennas on different APs can yield more diverse fingerprints.

For each RP, we construct 80 CSI images with dimensions of 180×180×3. After constructing CSI images for all RPs, they are fed into a CNN for training. The structure of the CNN is shown in Fig.3.

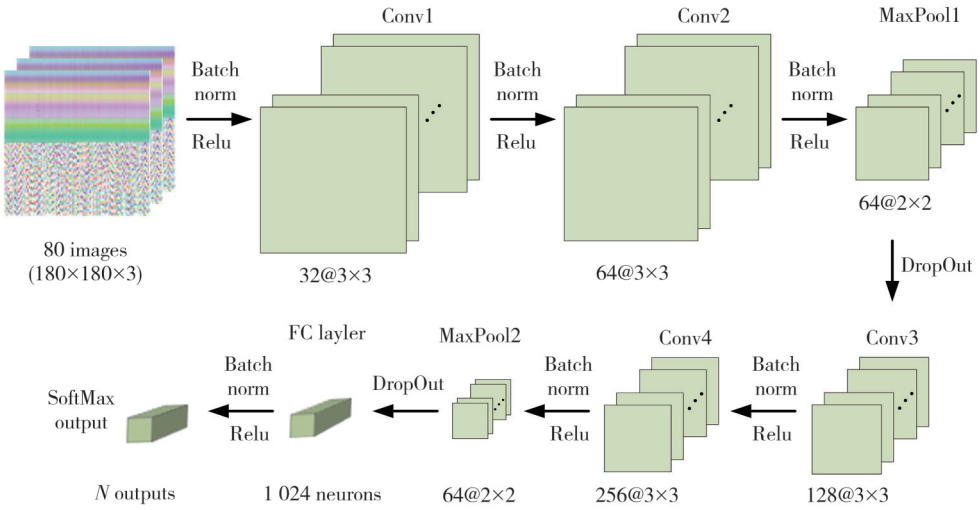


Fig. 3 Structure of CNN

The CNN consists of four convolutional layers, two max-pooling layers, one fully connected layer, and one output layer. Considering the similarity between images from different RPs, four convolutional layers with different sizes are employed to extract the features of images. Because the image size is relatively small, a stride of 1 is set for the convolutional kernels to ensure that the fully connected layer has sufficient input features. Two max-pooling layers are employed to capture key features in the images, which makes the features more robust. After each pooling, neurons are randomly dropped out with a dropout rate of 0.25, diminishing reliance on specific features. The fully connected layer maps the learned features to labels, which means it performs a linear transformation from one feature space to another feature space. The output layer employs the softmax function to transform the model's output into N probabilities, where N is the number of RPs. The training continues until the cross-entropy loss between adjacent iterations drops below a threshold, and the model is saved. Furthermore, stochastic gradient descent (SGD) is used to update the weights to optimize the training process, and batch normalization is introduced to accelerate error convergence. The parameters of the CNN are shown in Table 1.

Table 1 CNN parameter

Layer	Activation function	Kernel			Output
		Size	Count	Stride	
CSI Image					(180,180,3)
Conv1	ReLu	3×3	32	1	(178,178,32)
Conv2	ReLu	3×3	64	1	(176,176,64)
MaxPool1	ReLu	2×2	64	2	(88,88,64)
Conv3		3×3	128	1	(86,86,128)
Conv4	ReLu	3×3	256	1	(84,84,256)
MaxPool2	ReLu	2×2	256	2	(42,42,256)
Full connection					(1,1,1 024)
SoftMax output	ReLu				(1,1,N)

2.3 On-line phase

In the on-line phase, the first step is to collect the target device's CSI data. Subsequently, the image construction method is employed to convert CSI data into images. Finally, the constructed images are fed into the well-trained CNN model. The model's output is the probability values of the target device being located at each RP position, but in reality, the target device's position can be anywhere within the localization area. Therefore, select the top n outputs with the highest probability values and then use the weighted centroid algorithm to estimate the final position of the target device. The weighted centroid algorithm is

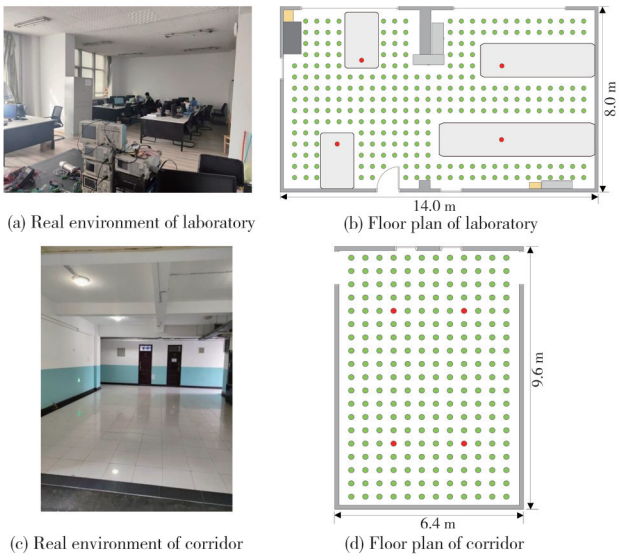
$$\bar{L}=\frac{\sum_{i=1}^{n}L_iP_i}{\sum_{i=1}^{n}P_i}, \tag{8}$$

where $\bar{L}$ is the estimated position coordinates, P_i is the probability value of the i-th location, and L_i is the coordinates of the i-th RP.

3 Experimental results and analysis

3.1 Experimental configuration

The experiment is conducted in a laboratory and a corridor. The two environments exhibit significant differences in signal interference and multipath effects. The corridor, measuring 9.6 m×6.4 m, is set with 236 RPs (marked by green dots in Fig. 4) and 4 APs (marked by red dots in Fig.4).



(a) Real environment of laboratory (b) Floor plan of laboratory

(c) Real environment of corridor (d) Floor plan of corridor

Fig. 4 Experimental environment

It has a strong openness and extremely low signal interference. It could be considered as a line-of-sight (LOS) environment. The laboratory, measuring 14 m×8 m, is set with 257 RPs (marked by green dots in Fig.4) and 4 APs (marked by red dots in Fig.4). The laboratory contains various objects such as tables, chairs, laboratory equipment, and storage cabinets, resulting in significant multipath effects, making it a non-line- of-sight (NLOS) environment. The distance between all neighboring RPs is 0.5 m. The real environment and floor plan of these two environments are shown in Fig.4.

We use four single-antenna TP-Link routers as APs and a Dell laptop as the receiving device. The laptop has a three-antenna Intel 5300 wireless card and runs the Ubuntu 22.04.3 LTS operating system. During the experiment, the laptop is set to monitor mode, with the frequency band set to 5 GHz and the bandwidth set to 20 MHz. It receives CSI data packets at a 1 ms interval.

For each RP, we collect 3 s of CSI data every other week by selecting three of the four APs with the highest RSSI. We conduct nine data collection sessions, resulting in a training dataset of 81 000 data packets per RP. After completing the collection of the training dataset, we collect 1 s of CSI data on three specific dates (one day later, one week later, and one month later) to construct the test dataset. We refer to the combination of the training dataset and test dataset as Dataset1. Data collection on different dates is because of the dynamic nature of indoor environments, influenced by factors such as human traffic, device mobility, and other variables. This approach allows us to capture the dynamic environment, thereby enhancing the adaptability and robustness of the localization system. Furthermore, to ensure a more fair comparison with other localization methods that typically use CSI data packets collected at the same time as their dataset, we collect 30 s of CSI data for each RP (a total of 90 000 data packets) at the same time. We refer to these data packets as Dataset2, where the initial 27 s serves as the training dataset and the subsequent 3 s as the testing dataset. After collecting all data, we apply the proposed method to convert all CSI data from the training dataset into images for training a CNN model. Each RP selects a subset of data packets from the test dataset to construct test images, which are then input into a well-trained CNN model for position estimation. Localization performance is evaluated by comparing the estimated and real positions according to localization results based on the test dataset.

3.2 Performance comparison of localization methods

To evaluate the proposed method, we employ the average localization error and the cumulative distribution

function (CDF) of errors as comprehensive evaluation criteria. We compare our method with ConFi[16], CiFi[8], and CNNFi[20]. To ensure fair experimentation, we use the CSI datasets (Dataset1 and Dataset2) collected in this paper for all methods. Dataset1 was collected at different times, while Dataset2 was collected at the same time. Error cumulative distributions for the laboratory and corridor environments are shown in Figs. 5 — 8, respectively. Tables 2 — 5 present the average localization error, standard deviation, and execution time for the laboratory and corridor environments.

Table 5 Performance comparison of different localization methods in the corridor using Dataset2

Algorithm	Mean error/m	Standard deviation/m	Execution time/s
Proposed method	0.183 6	0.174 2	0.750 3
CiFi	0.639 3	0.356 0	0.535 5
ConFi	0.812 4	0.473 2	0.523 1
CNNFi	0.278 7	0.213 1	0.575 2

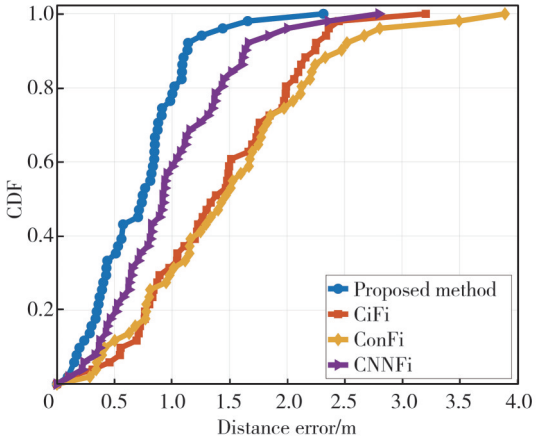


Fig. 5 CDF of different localization methods in laboratory using Dataset1

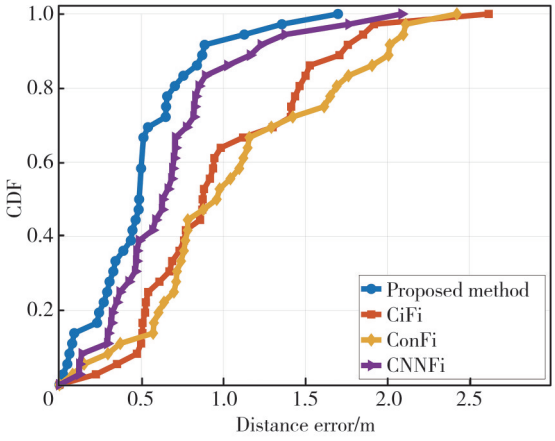


Fig. 6 CDF of different localization methods in corridor

The proposed method, ConFi, CiFi, and CNNFi use CSI as fingerprints and employ deep learning as the localization model. However, ConFi uses only the amplitude of the CSI as a fingerprint, while CiFi only employs the phase information of the CSI as a fingerprint. Neither method combines the amplitude and phase information of the AP's CSI, resulting in limited localization accuracy. In experiments using Dataset1 as the dataset, ConFi and CiFi achieve an average localization error of only 1.48 m and 1.39 m in laboratory environments and 1.07 m and 0.99 m in corridor environments, respectively. In experiments using Dataset2 as the dataset, ConFi and CiFi achieve an average localization error of only 0.97 m and 0.91 m in laboratory environments and 0.81 m and 0.64 m in corridor environments, respectively.

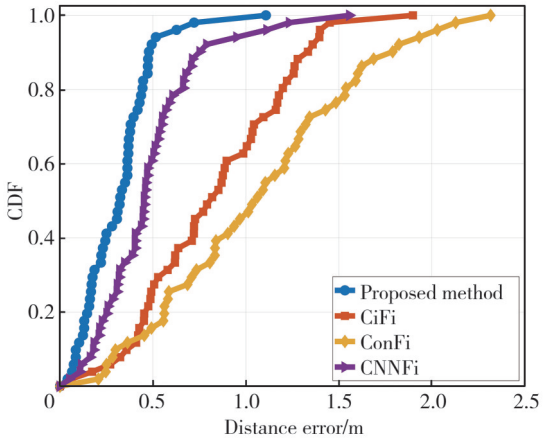


Fig. 7 CDF of different localization methods in laboratory using Dataset2

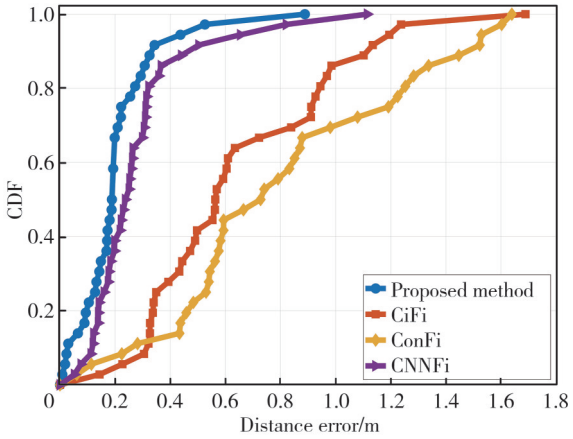


Fig. 8 CDF of different localization methods in corridor

Table 2 Performance comparison of different localization methods in laboratory using Dataset1

Algorithm	Mean error/m	Standard deviation/m	Execution time/s
Proposed method	0.726 0	0.424 0	0.751 9
CiFi	1.386 1	0.676 3	0.548 4
ConFi	1.481 7	0.817 9	0.529 6
CNNFi	0.987 2	0.566 1	0.483 7

Table 3 Performance comparison of different localization methods in corridor using Dataset1

Algorithm	Mean error/m	Standard deviation/m	Execution time/s
Proposed method	0.488 3	0.337 4	0.751 0
CiFi	0.991 4	0.552 1	0.536 2
ConFi	1.071 0	0.623 8	0.523 5
CNNFi	0.671 2	0.436 4	0.582 0

Table 4 Performance comparison of different localization methods in laboratory using Dataset2

Algorithm	Mean error/m	Standard deviation/m	Execution time/s
Proposed method	0.318 0	0.193 7	0.752 2
CiFi	0.909 6	0.443 7	0.543 8
ConFi	0.971 8	0.536 6	0.527 1
CNNFi	0.477 1	0.288 9	0.485 0

In contrast, the proposed method and CNNFi combine the amplitude and phase information of CSI to construct fingerprints, significantly enhancing localization accuracy. In experiments using Dataset1 as the dataset, the proposed method and CNNFi achieve average localization errors of 0.73 m and 0.98 m in laboratory environments and 0.49 m and 0.67 m in corridor environments, respectively. In experiments using Dataset2 as the dataset, the proposed method and CNNFi achieve average localization errors of 0.32 m and 0.48 m in laboratory environments and 0.19 m and 0.28 m in corridor environments, respectively. Compared to CNNFi, the proposed method improves localization accuracy by 0.25 m and 0.18 m in laboratory and corridor environments with Dataset1 and by 0.16 m and 0.1 m in laboratory and corridor environments with Dataset2, respectively.

There are three key reasons for the superiority of the proposed method over CNNFi. Firstly, the proposed method constructs a novel and more comprehensive dataset (Dataset1). Dataset1 covers a larger indoor environment having more complex signal propagation paths and reflection scenarios. The range interval of Dataset1 (0.5 m) is smaller than that of CNNFi (1.2 m). This increases the density and diversity of the data, thereby enhancing the resolution of the positioning system. Furthermore, the data in the CNNFi dataset is from the same period, while the data in Dataset1 spans different dates. This enables the reflection of the impact of environmental changes on the channel, thereby providing a more accurate description of the characteristics of this channel. Secondly, the proposed method improves the CNN model by introducing two maximum pooling layers to capture key image features. Furthermore, we employ Dropout following pooling to eliminate neurons randomly, thus enhancing model robustness, reducing reliance on specific features, and effectively preventing overfitting. Thirdly, the proposed method uses multiple APs for collecting CSI data, providing more fingerprint information. The CSI data from different APs can complement and correct each other, reducing errors and uncertainties. Moreover, removing APs with the lowest RSSI can improve CSI data quality. Because the proposed method uses CSI data from multiple APs for localization, the size of CSI images is relatively large, and the CNN model includes two additional max-pooling layers, resulting in increased processing time compared to other localization methods. However, it remains at approximately 0.75 s, meeting the real-time requirements of most indoor localization applications.

3.3 Impact of image construction methods on localization performance

To evaluate the impact of image construction methods on localization performance, we compare the proposed method, which combines amplitude and phase information of CSI, with two methods: one using only amplitude (referred to as "onlyAmplitude") and the other using only phase (referred to as "onlyPhase"). Images constructed using onlyAmplitude and onlyPhase are equivalent to the upper and lower halves of the image constructed by the proposed method, respectively. Furthermore, all experimental conditions remain consistent when comparing these three methods. Fig. 9 shows the average localization error using different image construction methods.

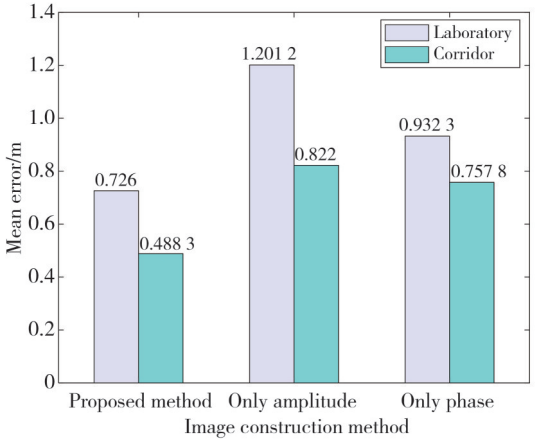


Fig. 9 Mean error of different image construction methods

The proposed method achieves average errors of 0.726 0 m and 0.488 3 m in laboratory and corridor environments, respectively. The localization accuracy of the proposed method surpasses onlyAmplitude and onlyPhase, confirming the effectiveness of combining amplitude and phase information.

3.4 Impact of number of training images on localization performance

To evaluate the impact of number of training images per RP on localization performance, we conduct training with different numbers of images (20, 40, 60, 80, 100,

120). Fig. 10 shows the average localization error using different numbers of images for training.

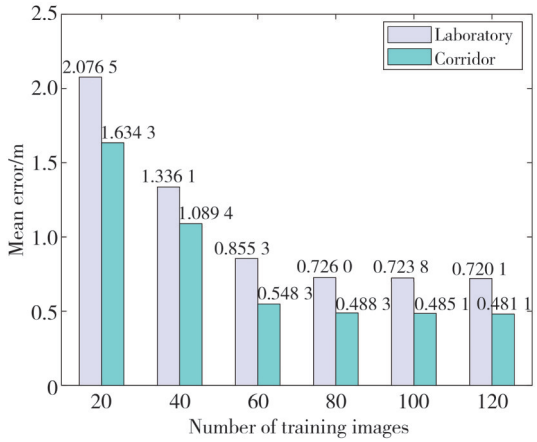


Fig. 10 Mean error of different numbers of training images

As shown in Fig. 10, an increase in the number of training images leads to a consistent reduction in localization error. When the number of training images is 20, the average errors are 2.076 5 m and 1.634 3 m in the laboratory and the corridor, respectively. However, with the number of training images raises to 80, the average errors decrease to 0.726 0 m and 0.488 3 m in the laboratory and corridor environments, respectively. Further increasing the training image number to 100 and 120 yields only marginal reductions in localization error compared to the 80. Consequently, the optimal number of training images is 80 to balance training time and localization accuracy.

4 Conclusions

This study proposed an indoor localization method that used CSI images from selected multi-APs. By selecting three APs based on RSSI for CSI data collection, we mitigated the impact of APs with poor signal quality on the localization results. For each RP, a novel three-channel CSI image was constructed by combining the amplitude and phase differences information from the different antennas of selected three APs. This image captured different signal propagation paths and multipath effects, providing richer fingerprint information for localization. Experimental results demonstrated that the proposed method achieved lower average localization errors in laboratory and corridor environments than CNNFi, ConFi, and CiFi.

Acknowledgement

This work was supported by Lanzhou Science and Technology Plan Project (No. 2023-3-104); Gansu Province Higher Education Industry Support Plan Project (No.2023CYZC-40); Gansu Province Excellent Graduate "Innovation Star" Program (No.2023CXZX-546)

Declaration of conflicting interests

The authors have no conflict of interests related to this publication.

References

[1] ALTıNPıNAR O V, SEZER V. A novel indoor localization algorithm based on a modified EKF using virtual dynamic point landmarks for 2D grid maps. Robotics and Autonomous Systems, 2023, 170: 104546.

[2] BI J X, ZHAO M Q, YAO G B, et al. PSOSVRPos: WiFi indoor positioning using SVR optimized by PSO. Expert Systems with Applications, 2023, 222: 119778.

[3] FARAHSARI P S, FARAHZADI A, REZAZADEH J, et al. A survey on indoor positioning systems for IoT-based applications. IEEE Internet of Things Journal, 2022, 9(10): 7680-7699.

[4] WANG Z H, XU Y Y. A Survey on privacy protection for indoor localization. Journal on Communications, 2023, 44(9): 188-204.

[5] LI Y C, YANG D X, GAO X, et al. Location and path planning in underground parking lot for intelligent vehicles. Optics and Precision Engineering, 2023, 31(5): 757-766.

[6] JIANG R, YU Y, XU Y Y, et al. Improved Kalman filter indoor positioning algorithm based on CHAN. Journal on Communications, 2023, 44(2): 136-147.

[7] YU M, YAO S Y, WU X, et al. Research on a Wi-Fi RSSI calibration algorithm based on WOA-BPNN for indoor positioning. Applied Sciences, 2022, 12(14): 7151.

[8] ALTAF KHATTAK S B, FAWAD, NASRALLA M M, et al. WLAN RSS-based fingerprinting for indoor localization: a machine learning inspired bag-of-features approach. Sensors, 2022, 22(14): 5236.

[9] ZHOU R, YANG Y X, CHEN P C. An RSS transform: based WKNN for indoor positioning. Sensors, 2021, 21(17): 5685.

[10] YANG X L, LI X Y, ZHOU M, et al. Multi-dimensional fuzzy mapping for AP optimization based WLAN indoor localization. Acta Electronica Sinica, 2022, 50(8): 1875-1884.

[11] LI Y B, SUN X. A highly robust indoor location algorithm using WiFi channel state information based on transfer learning reinforcement. Journal of Electronics & Information Technology, 2023, 45(10): 3657-3666.

[12] TONG X Y, ZHENG D C, GE W P, et al. Performance prediction for WiFi CSI localization system based on phased array. Journal of Software, 2023, 34(11): 5355-5375.

[13] KIM M. Graph-based machine learning for practical indoor localization. IEEE Sensors Letters, 2022, 6(12): 5501804.

[14] DING J Y, WANG Y, SI H Y, et al. Three-dimensional indoor localization and tracking for mobile target based on WiFi

sensing. IEEE Internet of Things Journal, 2022, 9(21): 21687-21701.
[15] DU L F, TIAN X Y, ZHANG L H, et al. Device-free indoor localization based on multidimensional CSI features classification. IEEE Access, 2023, 11: 32548-32563.
[16] CHEN H, ZHANG Y F, LI W, et al. ConFi: convolutional neural networks based indoor Wi-Fi localization using channel state information. IEEE Access, 2017, 5: 18066-18074.
[17] ZHU X Q, QU W Y, ZHOU X B, et al. Intelligent fingerprint-based localization scheme using CSI images for Internet of Things. IEEE Transactions on Network Science and Engineering, 2022, 9(4): 2378-2391.
[18] WANG X Y, WANG X Y, MAO S W. Deep convolutional neural networks for indoor localization with CSI images. IEEE Transactions on Network Science and Engineering, 2020, 7(1): 316-327.
[19] LI H H, ZENG X S, LI Y Z, et al. Convolutional neural networks based indoor Wi-Fi localization with a novel kind of CSI images. China Communications, 2019, 16(9): 250-260.
[20] LIU S, WANG X D, WU N. A CNN-based CSI fingerprint indoor localization method. Chinese Journal of Engineering, 2021, 43(11): 1512-1521.

来自选定多AP的信道状态信息图像室内定位

龙　良，王小鹏*，王　江，李　岗

兰州交通大学 电子与信息工程学院，甘肃 兰州 730070

摘　要： 为了提高使用信道状态信息（Channel state information，CSI）图像的室内定位方法的精度，提出了一种使用来自选定多个接入点（Access point，AP）的CSI图像的定位方法。该方法分为离线阶段和在线阶段。在离线阶段，根据接收信号强度指示（Received signal strength indication，RSSI）从定位区域的四个AP中选择三个。然后，使用商用Intel 5300网卡从三个选定的AP收集CSI数据。通过联合来自不同天线的振幅信息和相邻天线之间的相位差信息，为每个选定的AP构建一个单通道子图像。这些子图像合并为三通道的RGB图像，然后将其输入至卷积神经网络（Convolutional neural network，CNN）进行训练，训练完成后保存CNN模型。在线阶段收集来自目标设备的CSI数据，使用与离线阶段相同的步骤将其转换为图像，并输入至训练好的CNN模型。最后，使用基于模型输出概率值的加权质心算法估计目标设备的真实位置。提出方法在室内环境中使用两个数据集进行了验证，取得了良好的定位精度。

关键词： WiFi室内定位；多接入点指纹；信道状态信息图像；卷积神经网络；指纹定位；加权质心算法

引用格式： LONG Liang, WANG Xiaopeng, WANG Jiang, et al. Indoor localization with channel state information images from selected multiple access points. Journal of Measurement Science and Instrumentation, 2025, 16(4): 569-577. DOI: 10.62756/jmsi.1674-8042.2025055