NAS-FL:Fingerprint Localization Method Based on Automatically Designed Neural Network Architecture

Wen Liu[1,*], Haoyue Jiang[1,*], Ran Li[1] and Zhongliang Deng[1]

[1]School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract

With the development of deep learning technology, it has been widely used in indoor fingerprint-based localization. Existing indoor fingerprint localization methods based on deep learning carefully design the network structure of the model by manual means, and are used for location estimation in different environments. However, the spatial structure of different environments makes the signal propagation characteristics, location information and signal intensity distribution different, resulting in the universality of these methods when applied to different positioning scenarios. In this paper, a fingerprint localization system based on automatic design neural network structure is proposed, and the optimal network structure is carried out based on the two-layer optimization of micro-network architecture search algorithm. It can better extract the high-dimensional position information of the signal for location estimation of different spatial structure scenes. The model defines the unit module and simplifies the search of the whole network structure to the optimal unit structure. We conducted experiments in laboratory and library scenarios, which validated NAS-FL's superior performance in positioning accuracy and stability. The results show that this method can effectively improve the positioning accuracy and enhance the universality of the model in different positioning scenarios.

Keywords

Wireless signal, Automatic machine learning, Indoor localization, Fingerprint localization

1. Introduction

The complex building structure of indoor environment makes the signal face the problems of multipath propagation, non-line-of-sight transmission, signal attenuation and so on. The movement of people and ever-changing signal interference makes dynamic high-precision indoor positioning challenging[1]. The fingerprint location method based on WiFi technology makes full use of the widely deployed wireless network and environmental characteristic information, and can provide low-cost and low-complexity high-precision indoor positioning solutions, becoming the leading choice for indoor positioning[2]. With the popularity of deep learning applications, convolutional neural networks (CNN), graph neural networks (GNN), and long short-term memory artificial neural networks (LSTM) have been widely used in indoor fingerprinting localization. Traditional fingerprint localization methods based on deep learning carefully design the network structure of the model by manual means, and apply the model to sample location estimation in different environments. However, irregular spatial changes in different environments lead to different signal propagation characteristics, location information and signal strength distribution, resulting in some obvious inadaptability of these methods when applied to different spatial scenes.

Indoor positioning systems need to adapt to changes such as personnel movement, equipment switching, and building structure changes to maintain accuracy, and it has become a growing trend to design adaptive network models for different indoor environments. Neural architecture search (NAS)[3] is a kind of automated architecture engineering, which turns the process of adjusting neural networks according to experience into a process of automatically performing tasks to find more suitable architectures. The optimal structure of neural networks is designed for different data sets through

Proceedings of the Work-in-Progress Papers at the 14th International Conference on Indoor Positioning and Indoor Navigation(IPIN-WiP 2024), October 14 - 17, 2024, Hong Kong, China

*Corresponding author.

liuwen@bupt.edu.cn (W. Liu); Jiang_hy120@163.com (H. Jiang); lrliran1010@163.com (R. Li); dengzhl@bupt.edu.cn (Z. Deng)

0000-0002-6450-1969 (W. Liu); 0009-0001-8402-1516 (H. Jiang)

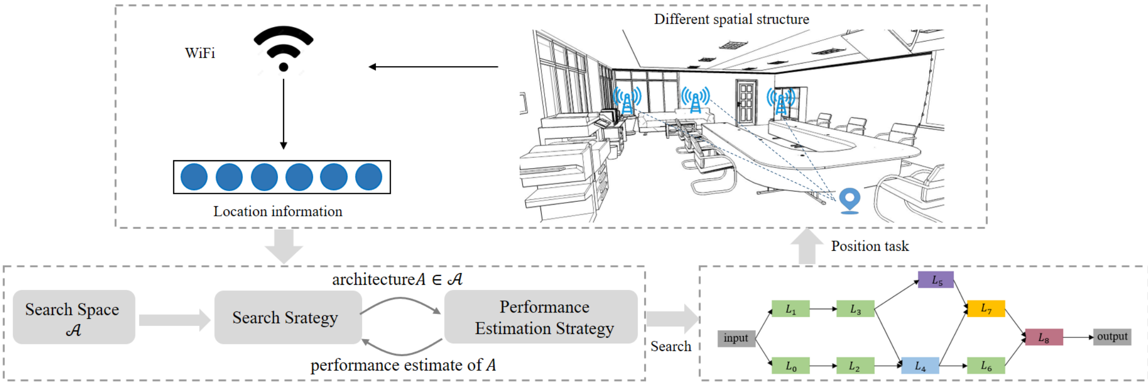


Figure 1: NAS-FL system overall frame diagram

search algorithms, and the automatically designed neural network architecture is better able to capture deeper feature information of different data sets. This improves the situation that the traditional method needs a lot of experiments and manual intervention to design the network structure, adjust the parameter information and select the optimization algorithm, but the manual design of the network structure has poor generalization performance on some data sets. Today, NAS methods are used in image classification, object detection, or semantic segmentation have outperformed manually designed architectures for some tasks, achieving highly competitive performance.

Inspired by the Neural Architecture Search (NAS) framework, this paper proposes an automatic design of neural network architecture for fingerprint localization: NAS-FL, which designs SOTA models for different indoor environments through search algorithms, and the automatically designed models can extract high-dimensional spatial location information in the environment, so that the localisation system has good generalisation ability under different spatial structures, and realises high accuracy and high robustness under the complex and variable indoor environment to achieve highly accurate and robust indoor positioning in complex and variable indoor environments. The system is shown in Fig. 1. The methodology and contributions of this work can be summarised as follows:

- First, define a rich search space, which simplifies the search of the entire network structure to the search of the optimal unit structure.
- Secondly, design a microizable network architecture search algorithm based on two-layer optimization. The algorithm adjusts the network architecture by gradient information, and can find a better network structure in less iterative steps.
- Finally, experiments on two indoor fingerprint data sets with different spatial structures show that the proposed method achieves excellent localization performance in different environments.

2. Related Work

Fingerprint localization: Deep learning is an effective technique for feature extraction and fingerprint matching. The fingerprint location algorithm based on deep learning can extract higher-order features from the original data and find the function between the data and the location, reducing the need for traditional feature engineering. In 2017, Chen[4]proposed the ConFi system , which designed a deep convolutional neural network to train and classify CSI amplitude images. The model consists of three convolutional layers and two fully connected layers to form a five-layer CNN. It can exploit local correlations by sharing the same weight between neurons in adjacent layers, outperforming DeepFi. In 2019, Hsieh et al.[5] designed a deep neural network based on MLP and 1D-CNN, using multiple one-dimensional convolutional layers to process CSI amplitudes and RSSI. The experimental results show that the designed 1D-CNN model network structure has excellent performance in extracting features.In 2023, Zhang et al.[6] proposed a domain adversarial graph convolutional network model. It designed the GCNS network for extracting graphic-level embeddings, effectively capturing the topology of the data. The traditional artificial network model design method has reached a relatively high positioning accuracy in fingerprint positioning, but its positioning accuracy varies with different indoor space structure, and the adaptability of the model is low in different environments. Therefore, this paper hopes to find a network model with high universality to better capture the spatial location information of WiFi signals in different structural environments and improve the robustness of the model in different indoor environments.

Neural Network Framework Search : Neural network Architecture Search (NAS) is a method of automatically learning and designing network structures to find an efficient neural network architecture for a specific task and data set. There are three main approaches to NAS[7]: the first approach is frame search based on evolutionary algorithms, which are capable of optimizing both the structure and weight of the network. While this approach has great potential in finding high-performance architectures, it has high computing resource requirements. The second approach is a framework search based on reinforcement learning (RL), which abstracts the design process of a neural network into a series of actions and uses the accuracy of the model as a reward signal to guide the search. This approach has made significant progress in automating network design, but still requires significant computational resources and time to achieve efficient searching. The third method is gradient-based frame search, which parameterizes function selection in the search space to continuous variables and then uses effective gradient information for backpropagation to accelerate the process of model search and construction. This method has advantages in improving search efficiency, especially in the case of limited computing resources. The model proposed in this letter is relevant to this category, because our goal is to make the model need to be able to adapt to changes in the indoor environment, such as personnel movement, equipment switching, and building structural changes, and further improve its efficiency and effectiveness while ensuring high precision positioning.

3. Proposed Method

In this paper, a WiFi based automatic design neural network framework for fingerprint localization (NAS-FL) is proposed, aiming at high robustness and high accuracy localisation in different dynamic indoor environments. The system framework diagram is shown in Fig.1. Based on the WiFi signals collected under different environmental spatial structures, the search strategy selects an architecture A from the predefined search space $\mathcal{A}$, passes the architecture to the performance estimation strategy, and the performance estimation strategy returns the estimated performance of A to the search strategy. The search strategy fully exploits the spatial location information contained in the WiFi signals under different spatial structures, and trains the search strategy through the above feedback mechanism to obtain the optimal network framework for the scenario. The training is carried out under this network framework to finally complete the location estimation in this scenario.

3.1. Search Space

Our proposed NAS-FL adopts the Cell-Based Network Architecture search method, which is divided into two kinds of structures, Normal Cell and Reduction Cell, and will constitute a complete network by splicing after the search is completed respectively. The core method is shown in Fig. 2.

A cell is a directed acyclic graph consisting of an ordered sequence of N nodes. Each node $x^{(i)}$ is a latent representation (e.g.a feature mapping in a convolutional network), and each directed edge (i, j) is associated with some operation $o^{(i,j)}$ that transforms $x^{(i)}$. We assume that the cell has two input nodes and one output node. For normal cells, the input node is defined as the cell output of the first two layers. For reduction cells, the input is defined as the input of the current step and the state carried from the previous step. The output of the cell is obtained by applying an approximation operation to all intermediate nodes.

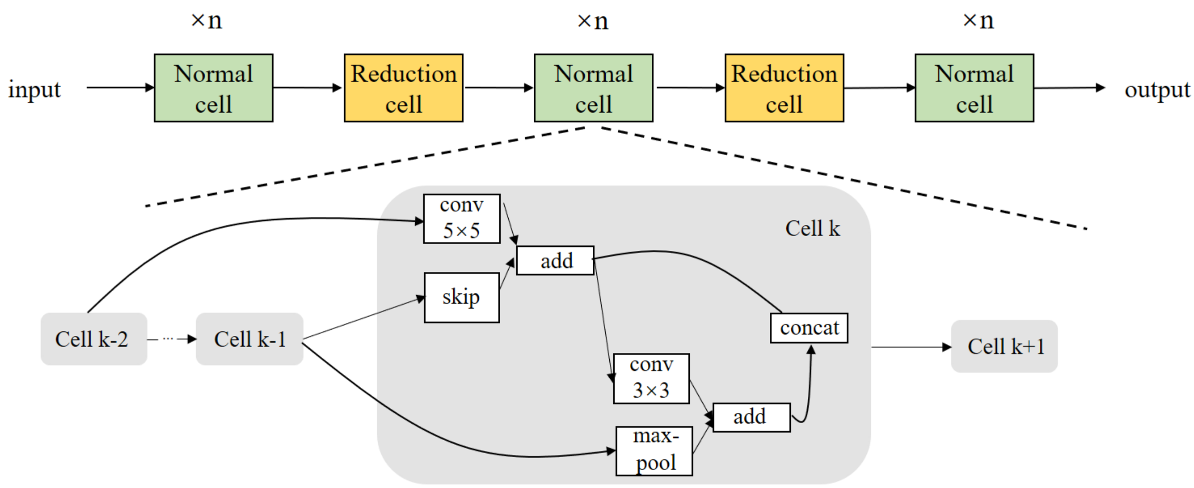


Figure 2: Diagram of Cell-Based network architecture

Table 1
Search Space

	Operations
O_n	max_pool 3x3, avg_pool_3x3, sep conv 3x3, sep conv 5x5, dil conv 3x3, dil conv 5x5 none
O_s	skip_connect, zero

Each intermediate node is calculated on the basis of all its predecessors:

$$x^{(j)} = \sum\nolimits_{i<j} o^{(i,j)} \left(x^{(i)} \right) \tag{1}$$

The task of learning a cell is reduced to learning the operations on its edges, and the possible operations between two nodes are selected in the search space. A well-designed search space is important for NAS methods. On the one hand, a good search space should be large enough and expressive enough to simulate various existing CNN models, thus ensuring that the performance of the models it searches is competitive. On the other hand, the search space should be small and compact enough to save computational resources. The expressiveness of a neural network model depends on the properties of different aggregation functions, so we design an expressive and simplified search space:

- Node Aggregators: We chose six node aggregators with good results according to the popular CNN models, as shown in Table 1. We use O_n to denote the node aggregators.
- Connection operation: We use O_s connection operation. zero denotes a special zero operation that represents no connection between two nodes.

3.2. Continuous relaxation and optimization

Let O be a set of candidate operations in O_n, where each operation denotes some function $O(\cdot)$ that will be applied to $x(i)$. In order to make the search space continuous, we relax the classification choice of a particular operation to the $softmax$ value of all possible operations:

$$o^{(i,j)}(x) = \sum\nolimits_{o \in O_n} \frac{\exp\left(\alpha_o^{(i,j)}\right)}{\sum\nolimits_{o' \in \mathcal{O}} \exp\left(\alpha_{o'}^{(i,j)}\right)} o(x) \tag{2}$$

where a pair of nodes (i,j) is parameterised by a vector $\alpha^{(i,j)}$ of dimension $|O_n|$. The task of architectural search is then reduced to learning a set of continuous variables $\alpha = \left\{\alpha^{(i,j)}\right\}$, as shown in Fig. 3.

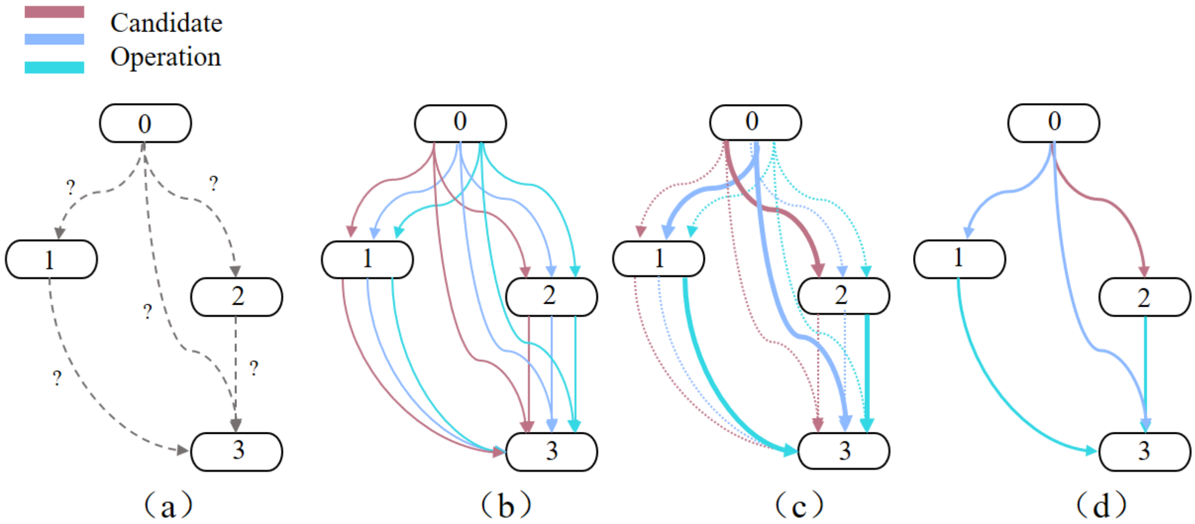


Figure 3: Overview of the search process: (a) The operation on the edge is initially unknown. (b) Continuously widen the search space by placing mixed candidate operations on each edge. (c) By solving the two-layer optimization problem, the mixed probability and the network weight are jointly optimized. (d) The final structure is derived from the learned mixing probabilities.

At the end of the search, by replacing each mixed operation $o^{(i,j)}(x)$ with the most probable operation, i.e:

$$o^{(i,j)} = \text{argmax}_{o \in O_n} \alpha_o^{(i,j)} \tag{3}$$

Once the training is complete, the edge with the highest probability is found from all the edges to form the discrete architecture.

3.3. Search Algorithm

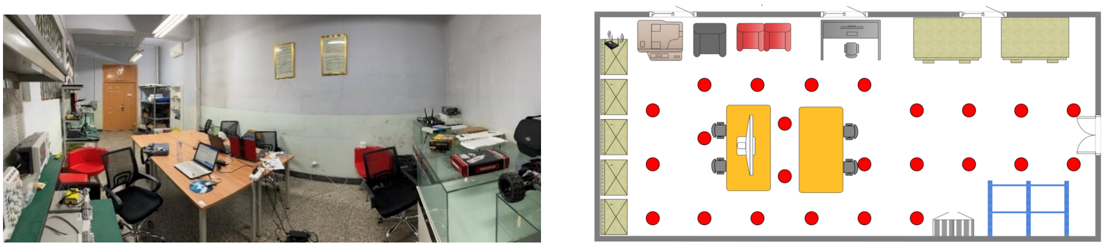
After relaxation, the structure α and the weights w (e.g.the weights of the convolutional filters) need to be learnt jointly for all hybrid operations. Similar to architectural search using RL or evolutionary algorithms, where the validation set performance is considered as a reward or fitness, the search algorithm in this paper aims to optimise the validation loss but using gradient descent. $\mathcal{L}_{\text{train}}$ and $\mathcal{L}_{\text{val}}$ denote the training loss and validation loss, respectively. Both losses depend not only on the structure α, but also on the weights in the network w. The goal of the architectural search is to find the validation loss that makes the validation loss $\mathcal{L}_{\text{val}}(w^*, \alpha^*)$, where the architecture-dependent weights w^* are obtained by minimising the training loss $w^* = \text{argmin}_w \mathcal{L}_{\text{train}}(w, \alpha^*)$.

The system views the search problem as a second-level optimisation problem with α as the upper-level variable and w as the lower-level variable:

$$\text{min}_\alpha \mathcal{L}_{\text{val}}\left(w^*(\alpha), \alpha\right) \text{ s.t. } w^*(\alpha) = \text{argmin}_w \mathcal{L}_{\text{train}}(w, \alpha) \tag{4}$$

3.4. Enhanced Optimization algorithm

In the two-layer optimisation algorithm for gradient, the need to wait for the sub-network training to converge after each search may result in the optimisation process failing to converge to a (locally) optimal solution, resulting in a sub-optimal network framework, so we optimise the search algorithm by using $\mathcal{L}_{\text{val}}(w, \alpha)$the idea of regularisation, using as a regular term, adding constraints to the above



(a) (b)

Figure 4: Schematic diagram of laboratory environment: (a)The real scene from the lab (b)Laboratory reference point layout

equation, where δ is a constant, to obtain a hybrid optimisation algorithm:

$$\begin{gathered} \omega^*(\alpha) = \arg\min_{\omega} \mathcal{L}_{\text{train}}(\omega, \alpha), \\ \min_{\alpha} \left(1-\lambda'\right) \mathcal{L}_{\text{train}}\left(\omega^*(\alpha), \alpha\right) + \lambda' \mathcal{L}(\omega^*(\alpha), \alpha) - \lambda' \delta, \\ 0 \leq \lambda \leq 1 \end{gathered} \tag{5}$$

Lagrange multipliers and normalisation are performed to finally obtain a first order mixed layer optimisation Loss:

$$\min_{\alpha,\omega} \left[\mathcal{L}_{\text{train}}\left(\omega^*(\alpha), \alpha\right) + \lambda \mathcal{L}_{\text{val}}\left(\omega^*(\alpha), \alpha\right)\right] \tag{6}$$

where is a non-negative regularisation parameter that balances the importance of training loss and validation loss. Thus, by taking into account the potential relationship between training loss and validation loss, our hybrid-level optimisation can alleviate the overfitting problem and search the architecture with higher accuracy than the two-level optimisation.

4. Experiments Validation

4.1. Comparison with Based Localization Mode

In order to verify the model robustness of the system under different spatial structures, we tested the localisation stability and accuracy of NAS-FL in two typical indoor scenarios, a small laboratory with a complex environment and a relatively empty large library.

Laboratory environment: The laboratory dataset consists of CSI data collected from WiFi signals in our small lab. The CSI data acquisition device consists of a transmitter and a receiver with three antennas. RP selection is a necessary part of fingerprint positioning, and too large or too small an interval will affect the effect of the experiment. We consider positioning accuracy and data acquisition effort to evenly select RP locations at appropriate intervals. We chose a distance of 0.8 m between adjacent RPS in the laboratory and set 24 reference points. This ensures good positioning accuracy and makes offline data collection workloads acceptable. The CSI of each reference point was collected for two different periods of time. These two periods each contain three different layouts of indoor obstacles. Each sample is composed of 30 consecutive CSI packets. The data collected in the first session was used to build an offline training set with 40 training samples per reference point. Data from the second session were used to form an online test set, again with 40 test samples collected at each reference point. All data is searched into the final data set. The specific distribution is shown in Figure 4.

Library environment: We use the open indoor positioning dataset-10.5281/zenodo.1066040 provided by scholars Mendoza-Silva G et al. as another experimental dataset.This dataset collects the RSSI data values of the third and fifth layers of a library over a period of 15 months and contains rich environmental characteristics. It has 448 unique access Points (aps) with MAC address identifiers. At 48 reference points (RPS), measurements of signal strength were collected. In order to ensure the accuracy of the

(a) (b)

Figure 5: Schematic diagram of library environment: (a)The real scene from the library (b)Library reference point layout

data, each RP is pre-assigned a serial number, and the data is collected in sequence according to the serial number. Six collections were made at each reference point, which reduced errors. The monthly data is divided into 1 training set and 5 test sets. The distribution of data collection points is shown in Figure 5.

4.2. Evaluation Metric

In order to comprehensively analyse the localisation performance of the model, we considered the following evaluation metrics during the evaluation process:localisation error, average localisation error, and standard deviation. The average positioning error and standard deviation reflect the positioning accuracy and stability of the fingerprint positioning model, respectively.

For test sample i, the localisation error represents the distance between the predicted coordinates and the actual coordinates, which is calculated as

$$\text{error}_i = \sqrt{(x_i - x_{\ i})^2 + (y_i - y_{\ i})^2} \tag{7}$$

where (x_i, y_i) are the true coordinates of test sample i and $(x_{\ i}, y_{\ i})$ are the coordinates predicted by the localisation model. For all test samples, the mean error and standard deviation of localisation were calculated, respectively, by Eq:

$$\text{Mean} = \frac{1}{N}\sum_{i=1}^{N} \text{error}_i, \text{Std} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (\text{error}_i - \text{Mean})^2}, \tag{8}$$

Here N is the total number of test samples. The error accumulation distribution function (CDF) curve is used to represent the positioning error distribution of the test sample.

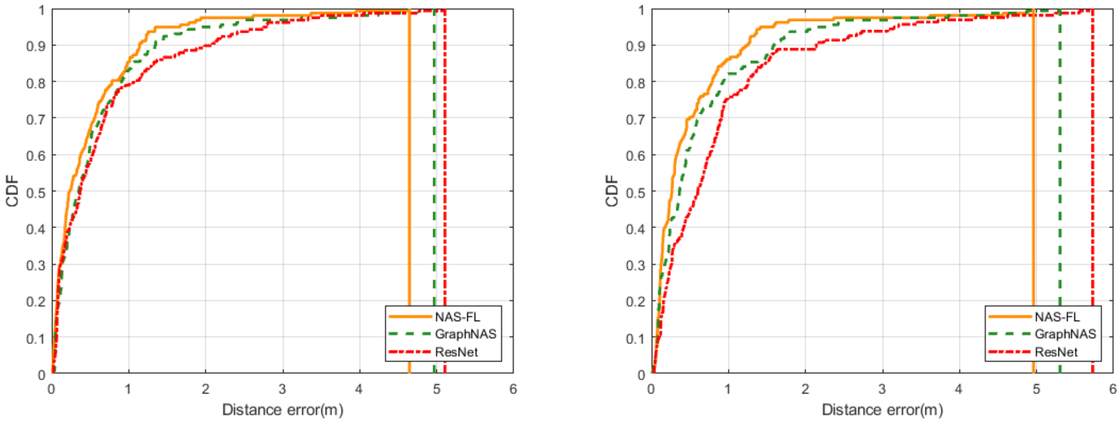
4.3. Experimental Results and Analysis

To validate the performance of our system in searching network architectures, we compared it to two other localization methods for manually designing network frameworks, including the ResNet model based on manually designed network architectures and the GraphNAS model based on graph structures for automatically searching network architectures.

The comparison of the localization results in Table 2 shows that the overall performance of the model obtained from our search calculation is better than the other two methods. Compared with the manually designed ResNet, ResNet's model for processing WiFi signals under different spatial structures remains unchanged, and the ability of a single CNN model to process CSI image features is weaker.GraphNAS can extract and integrate the location information contained in the antenna link amplitude and phase,

Table 2
Positioning performance in two spatial environments (Mean and Std)

Case	Laboratory		Library	
Method	Mean(m)	Std(m)	Mean(m)	Std(m)
ResNet	0.82	1.18	1.08	1.37
GraphNAS	0.68	1.14	0.89	1.26
NAS-FL	0.55	0.97	0.73	1.18



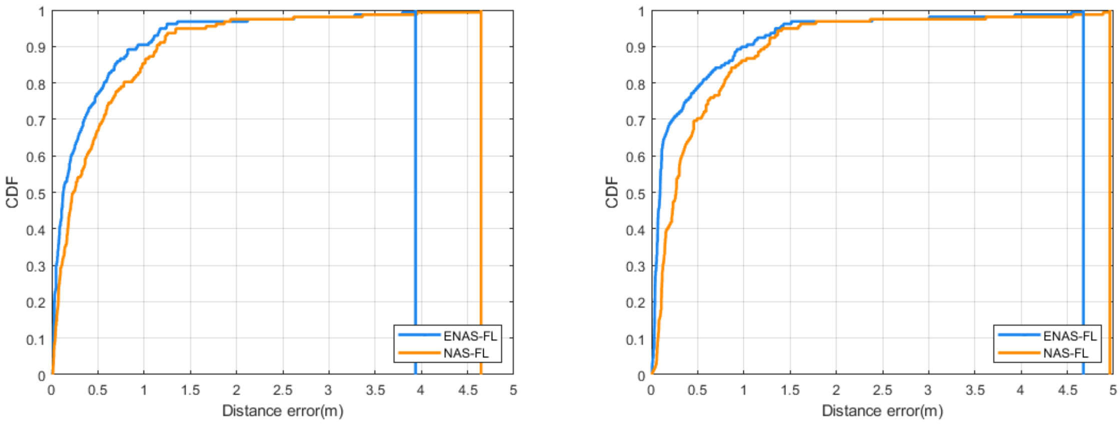
(a) (b)

Figure 6: CDF of NAS-FL and manual design model positioning errors in laboratory and library scenarios

and make better use of the spatial diversity characteristics of CSI, but the complex changes in the spatial scene and the unchanging model structure reduce the model's adaptability. According to the different spatial structures of CSI signals, NAS-FL automatically searches for the model structure that is compatible with the environment through the search algorithm, so that the feature information of CSI images can be better extracted and accurate positioning can be carried out.

Figure 6 (a) shows the cumulative positioning errors of NAS-FL, ResNet, and GraphNAS in the laboratory scenario. In the 0-2 m range, NAS-FL showed good positioning performance, with GraphNAS and ResNet slightly worse. The search algorithm in NAS-FL combined with the network structure of different scene spatial structure search, has a strong ability to extract the spatial location information of CSI image features, and effectively improves the generalization ability of the model. When the probability reaches 1.0, the positioning error of NAS-FL is 4.50 m, that of GraphNAS is 4.95m and that of ResNet is 5.15m. Figure 6 (b) shows the cumulative positioning error of NAS-FL, ResNet, and GraphNAS in the library scenario, which is slightly worse than the positioning error in the laboratory scenario. In general, both NAS-FL and GraphNAS have good positioning performance within 0-1m, and are better than ResNet. GraphNAS has initially performed well, but the trend has been erratic, which has created a lot of uncertainty about the positioning of practical applications. When the probability reaches 1.0, the positioning error is 4.92 m for NAS-FL, 5.36m for GraphNAS, and 5.79m for ResNet. NAS-FL has more stable performance, less positioning error, and stronger model robustness.

Figure 7(a) shows the cumulative positioning errors of NAS-FL and ENAS-FL in the laboratory scenario. In the 0-2 m range, both have good positioning ability, ENAS-FL is slightly better. The location error of NAS-FL is small, and the hybrid optimization search algorithm converges in the search process to obtain the local optimal solution, which makes the model more accurate. Figure 7(b) shows the cumulative positioning errors of NAS-FL and ENAS-FL in the laboratory scenario. In the 0-0.8m range, the positioning performance of both is very good. NAS-FL has more stable performance and less positioning error, and the hybrid optimization search network architecture effectively reduces the error fluctuation range.



(a) (b)

Figure 7: CDF of location error of two-layer optimization algorithm and enhanced optimization algorithm in laboratory and library scenarios

5. Conclusion

In this paper, we propose a fingerprint localization method based on automatically designed neural network architecture. The method can design an adaptive network model according to different spatial structures and fully mine the location information of each space. According to the WiFi signal in the input scene, the network model structure is searched based on the feedback mechanism of gradient two-layer optimization strategy. We conducted experiments in two different spatial structure environments with different signal characteristics, and NAS-FL showed good positioning accuracy and stability, which verified that the algorithm can improve the generalization ability of the positioning system in different positioning scenarios, and realize highly robust fingerprint positioning in complex indoor environments.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China under Grant No.62372049.

References

[1] W. Liu, Q. Cheng, Z. Deng, H. Chen, X. Fu, X. Zheng, S. Zheng, C. Chen, and S. Wang, “Survey on csi-based indoor positioning systems and recent advances,” in *2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN).* IEEE, 2019, pp. 1–8.

[2] S. Hayward and A. West, “A survey of indoor location technologies, techniques and applications in industry,” *Internet of Things*, p. 100608, 2022.

[3] P. Ren, Y. Xiao, X. Chang, P.-Y. Huang, Z. Li, X. Chen, and X. Wang, “A comprehensive survey of neural architecture search: Challenges and solutions,” *ACM Computing Surveys (CSUR)*, vol. 54, no. 4, 2021.

[4] H. Chen, Y. Zhang, W. Li, X. Tao, and P. Zhang, “Confi: Convolutional neural networks based indoor wi-fi localization using channel state information,” *Ieee Access*, vol. 5, pp. 18 066–18 074, 2017.

[5] C.-H. Hsieh, J.-Y. Chen, and B.-H. Nien, “Deep learning-based indoor localization using received signal strength and channel state information,” *IEEE access*, vol. 7, pp. 33 256–33 267, 2019.

[6] M. Zhang, Z. Fan, R. Shibasaki, and X. Song, “Domain adversarial graph convolutional network based on rssi and crowdsensing for indoor localization,” *IEEE Internet of Things Journal*, 2023.

[7] S. Liu, H. Zhang, and Y. Jin, “A survey on computationally efficient neural architecture search,” *Journal of Automation and Intelligence*, vol. 1, no. 1, p. 100002, 2022.