IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 7 March 2023, accepted 1 April 2023, date of publication 13 April 2023, date of current version 19 April 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3266874

TOPICAL REVIEW

Low-Cost Indoor Wireless Fingerprint Location Database Construction Methods: A Review

WEN LIU, YINGENG ZHANG, ZHONGLIANG DENG, (Senior Member, IEEE), AND HEYANG ZHOU
School of Electronics Engineering, Beijing University of Posts and Telecommunications, Beijing 100089, China
Corresponding author: Wen Liu (liuwen@bupt.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61871054.

ABSTRACT The fingerprint positioning has achieved remarkable results in indoor localization tasks, but the method usually relies on a large amount of fingerprint data to build a fingerprint database, and the amount and diversity of fingerprint data will directly affect the effectiveness of fingerprint positioning. Since fingerprint acquisition is limited and disturbed by space and time, it consumes a lot of labor and time costs to collect fingerprint data in the localization environment, and wireless fingerprint data is time-sensitive and environment-dependent, and changes in the localization environment will reduce the usability of the existing fingerprint database. The complex and repetitive fingerprint acquisition work seriously affects the feasibility of practical deployment of fingerprint positioning systems in the positioning environment. Therefore, the study of low-cost wireless fingerprint database construction methods has become an inevitable part of promoting the widespread deployment of indoor fingerprint positioning systems. In this paper, we introduce the traditional data augmentation-based approach and the advanced machine learning model-based approach, systematically presenting the underlying models and algorithms of both. The former reviews the application of two traditional data enhancement methods, namely channel propagation models and interpolation or regression, to the construction of low-cost wireless fingerprint databases, while the latter taps into techniques for reducing the cost of fingerprint database construction by combining generative adversarial networks and small-sample learning models with the indoor localization domain. Finally, we discuss the current challenges and future research directions for achieving high-performance indoor localization based on low-cost wireless fingerprint databases, and suggest some useful research guidelines.

INDEX TERMS Indoor fingerprint location, wireless sensor network, data augmentation, machine learning.

I. INTRODUCTION

As an important part of indoor location services, indoor positioning is an important foundation for applications such as indoor navigation, security monitoring, and emergency disaster relief [1], [2]. With the emergence of high-performance and highly generalized positioning requirements, among many indoor positioning technologies, fingerprint positioning has been favored by researchers due to its high positioning accuracy, low equipment cost and easy deployment [3], [4], [5], [6].

Fingerprint-based indoor localization systems generally rely on a wireless local area network (WLAN) deployed in the localization environment, which usually consists of two phases: offline phase and online phase [7]. The offline phase usually collects received signal strength (RSSI) or channel station information (CSI) at each reference point(RPs) as fingerprints to build a fingerprint database [8], [9], [10] and train a classification model for the online phase; the online phase collects real-time data and compares it with the fingerprint database to obtain predicted device location information for localization.

The localization effect of fingerprint localization system mainly depends on the matching degree between the fingerprint to be located and its real location fingerprint, and the diversity and quantity of fingerprint data collected in the offline phase will significantly affect the localization accuracy of fingerprint localization system [11], [12]. Properly

The associate editor coordinating the review of this manuscript and approving it for publication was Marco Martalo.

setting fingerprint points, collecting high-quality fingerprint data and correctly marking location information are the prerequisites for fingerprint-based localization methods to achieve high-precision localization results. Fingerprint localization accuracy usually increases with the number of fingerprint points, but as the distance between fingerprint points becomes smaller, the gain in localization accuracy due to the increase in the number of fingerprint points will slow down, and the localization accuracy will gradually improve with the increase in the number of fingerprint samples collected at each fingerprint point [13], and a larger fingerprint data set will produce more satisfactory results than a smaller data set [14], [15]. However, the workload to collect a complete offline fingerprint dataset in the field and to label it correctly will increase geometrically with the number of preset fingerprint points in the localization scene [16], and as coarse-grained information, the time sensitivity and environmental dependence of the wireless signal will cause the collected fingerprint data to fluctuate due to the multipath effect in indoor environments, in complex scenes with dense and changing environments, thus leading to significant differences between the offline fingerprint database and online data distribution [17], [18], which seriously affects the performance of fingerprint localization systems. Therefore, in order to achieve high-precision indoor positioning effects, a large amount of fingerprint data needs to be repeatedly collected many times at the same fingerprint point located under different spatial and temporal conditions, which greatly increases the complexity and labor cost of widely deployed indoor fingerprint positioning systems and seriously hinders the popularity and application of indoor fingerprint positioning systems [19].

Spatial and temporal constraints and perturbations on fingerprint acquisition make it difficult to collect sufficient and uniform amounts of fingerprint data under different spatio-temporal conditions and densely deployed fingerprint points, leading to many challenges in achieving highly accurate and robust fingerprint localization methods [20], [21]: a) Densely deploying fingerprint points and collecting large amounts of marker data is both time-consuming and labor-intensive. b) Dynamic changes in the localization environment can lead to differences between the online fingerprint distribution and the fingerprint database, requiring multiple repetitions of fingerprint data acquisition under multiple spatial and temporal conditions. c) Fingerprint data are environment-specific, and when the localization environment is changed, data acquisition needs to be repeated in the new environment. In response to the above challenges, studying how to build a wireless fingerprint database at low cost to achieve high accuracy and robust localization is an important way to reduce the labor and time costs required for fingerprint data collection for indoor positioning and to improve the feasibility of practical deployment of fingerprint positioning systems in multiple positioning scenarios.

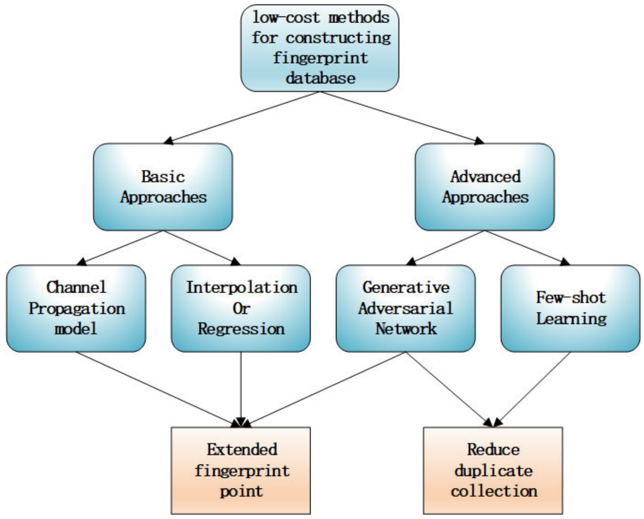


FIGURE 1. A taxonomy of Efficient use of fingerprint data methods.

In this paper, we aim to provide theoretical support and directional guidance to address the above challenges by outlining and summarizing existing methods for building wireless fingerprint databases at low cost. As shown in Figure 1, we review the classical channel propagation models, interpolation or regression to achieve fingerprint data augmentation, which use fingerprint data collected from actual fingerprint points to generate virtual fingerprint point data for fingerprint database expansion. In addition, we introduce the application of two advanced machine learning models, generative adversarial networks and few-shot learning. We will also discuss the future direction of studying how to build a wireless fingerprint database at low cost to support the deployment of high accuracy and highly robust localization systems. The rest of this paper is organized as follows. Section II details the application of two traditional data enhancement methods, channel propagation model and interpolation or regression, to expand fingerprint data points in fingerprint localization. Section III details the application of two advanced machine learning models, generative adversarial network coupling and small sample learning, to improve fingerprint data utilization in fingerprint localization. Section IV concludes the whole paper.

II. BASIC DATA AUGMENTATION METHODS

Fingerprint positioning methods are favored by researchers in the field of indoor localization; however, fingerprint localization relies heavily on the amount of fingerprint data and the density of fingerprint points to achieve high robust performance, and therefore deploys as many fingerprint points as possible in a large scale during the fingerprint data acquisition phase [4]. Unfortunately, in practice, some locations in the localization scenario may be inaccessible or restricted by obstacles, making it difficult to deploy fingerprint points (RPS) densely and regularly throughout the region [22], and the time and labor cost required for acquisition increases with the size of the environment, the number of fingerprint points, and the amount of data collected per fingerprint point [23]. Although reducing the density of fingerprint point deployment and the number of fingerprint data collected can reduce costs, it can lead to rapid performance degradation of the fingerprint localization system and significant reduction in localization accuracy as the number of fingerprint points decreases beyond a threshold [13]. Some researchers have investigated the generation of virtual fingerprints by two traditional means of data enhancement: 1) locally weighted interpolation of the initial measurement set and 2) indoor channel propagation modeling by computing scene-specific propagation parameters. In these studies fingerprint points are divided into real and virtual points, where real points are the points where fingerprint data are collected in the offline phase in fingerprint localization methods, while virtual points are fingerprint points and fingerprint data automatically generated based on fingerprint data from the above two methods and real points. These studies artificially increase the number of "fingerprint points" without adding additional acquisition costs. Using traditional data augmentation techniques to generate virtual fingerprint data is an effective way to build fingerprint database at low cost, which can enrich fingerprint database information, achieve intensive fingerprint coverage degree, reduce the cost of intensive fingerprint acquisition, and provide an effective way to solve the difficult problem that it is costly and difficult to collect fingerprint data from densely deployed fingerprint points. In this section, we present the positioning methods based on 1) channel propagation model; 2) interpolation or regression, two traditional data enhancement means to achieve low-cost construction of wireless fingerprint database.

A. METHOD BASED ON CHANNEL PROPAGATION MODEL

In indoor environments, wireless signal transmission is affected by factors such as multipath interference, shadowing effect, power fading, and transmission delay [24], [25]. It has been shown that various obstacles such as walls and homes existing between wireless signal transceivers in the localization scenario can cause multipath fading and shadowing effects by absorbing, reflecting, scattering, and diffracting resulting in wireless signal power attenuation [26], therefore, the complexity and limitations of indoor environment and obstacles should be fully considered when modeling channel propagation, combining elements such as path fading, scattering, refraction, and reflection for modeling. Typical channel propagation models include

1) INDOOR GENERAL PROPAGATION MODEL

$$L(d) = L + 10 \times n_{sf} \times \log_{10}(d) + FAF \tag{1}$$

The model can be used in a typical indoor environment, where L denotes the path loss per unit distance of wireless signal transmission in free space, n_{sf} denotes the loss factor of the same layer, FAF denotes the additional value of loss due to multipath, and d denotes the distance between the transmitter and the receiver. This model does not consider the obstacle distribution in the actual space, only the environmental path loss information is needed, and the average path loss and the related shadow fading statistics are used to characterize the indoor path loss.

2) CHAN PROPAGATION MODEL

$$L = L_p + +L_W \tag{2}$$

The model is applicable to indoor wireless signal strength prediction, where L denotes the path loss of wireless signal as it propagates indoors, L_p denotes the path loss of the wireless signal as it propagates in free space, and L_W denotes the loss of wireless signal as it penetrates obstacles such as walls.

3) LOSS FACTOR PROPAGATION MODEL

$$L(d) = L + 20\log(\mathrm{d}) + \alpha d + FAF \tag{3}$$

For a complex indoor environment, the indoor path loss is equal to the free space loss plus the loss factor and increases exponentially with the distance between the transmitter and the receiver. Where L is the path loss per unit distance of wireless signal transmission in free space, α is the attenuation constant of the channel in unit dB/m, d is the distance between the transmitter and the receiver, and FAF denotes the loss additive caused by multipath.

4) KEENAN-MOTELY PROPAGATION MODEL

$$L = L(d_0) + 20\log\left(\frac{d}{d_0}\right) + \sum_{j=1}^{j} N_{wj}L_{wj} + \sum_{i=1}^{i} N_{Fi}L_{Fi} \tag{4}$$

The model is suitable for simulating complex indoor path loss with multiple walls on multiple floors, where d is the distance between transmitter and receiver, N_{wj} and N_{Fi} denotes the number of signals passing through different types of walls and floors, respectively, L_{wj} and L_{Fi} correspond to their loss factors, and j and i denotes the number of types of walls and floors, respectively.

The fingerprint positioning method based on channel propagation model can effectively reduce the cost required to collect fingerprint data to build a complete data set by using a small amount of measured real fingerprint data to construct a wireless signal propagation model based on full consideration of the building information and environmental layout of indoor scenes.

As shown in Table 1, in order to reduce the cost and shorten the fingerprint data collection time in the offline data collection phase of fingerprint localization, some studies have used channel propagation models to generate fingerprint data information to expand the fingerprint database and enhance the feasibility of fingerprint localization system deployment. In 2000, [13] proposed to generate fingerprint data using a path loss model to train log distance path using the actual measured fingerprint data at RPS loss model parameters to

TABLE 1. Low-cost indoor location method based on channel propagation model.

Ref	Year	Parameter and Technology	Position error(m)	interpolation errors(dBm)	Scene (m^2)
[13]	2000	RSSI, NIC	4.3	9.82	43.5*22.5
[36]	2006	RSSI, WIFI	1.7	×	16*10
[37]	2010	RSSI, WIFI	3	×	27*18
[29]	2013	RSSI, WIFI	3.4	3.7	15*60
[34]	2016	RSSI, WIFI	1.25	×	6*6
[35]	2017	RSSI, WIFI	3.42	×	20*30
[32]	2017	RSSI, WIFI	3.91	4.31	85*76
[30]	2019	RSSI, WIFI	2.95	3	211*2.4
[31]	2019	RSSI, WIFI	1.44	<3	23*11

× : not covered in the article

generate RSSI fingerprint data at locations where RPS is not set in the localization scenario, and thus construct a complete fingerprint database. RSSI fingerprint data with an average error of 9.82 *dBm* was generated in a localization environment of 43.5m × 22.5m. Due to the complexity of indoor environment, the propagation model has difficulty in accurately predicting signal fading and multipath effects, resulting in poor localization based on generated fingerprints compared to measured fingerprints [27], [28]. In order to improve the accuracy of fingerprint generation, [29] expanded the measurement database using a path loss model that considers multi-wall fading and used the number of walls spaced between access points(APs) and RPs as one of the parameters of generated fingerprint data, and the average error between the generated virtual RSSI and the real value was 3.7*dbm*. achieving a localization accuracy of 3.4m in a complex indoor area of 15m × 60m. However, the method requires a priori knowledge of the structure and layout of the localization environment as support. Based on the area-based approach, [31] proposed a method that does not require the environment layout and the number of walls as a priori knowledge, sets different path loss parameters for each area, and generates RSSI fingerprint data for different areas by combining the structural characteristics of indoor areas and the environmental impact. In the complex positioning environment of 23m × 11m, the virtual RSSI fingerprint with average error less than 3 *dBm* were generated,and a localization effect with an average error of 1.44m was achieved.In [32], an adaptive model parameter configuration algorithm was developed to determine the optimal parameters of the channel propagation model with limited fingerprint data, and a probabilistic method was used to calculate the accuracy of the estimated RSSI prediction values in the localization environment.In the localization scenario of 85m × 76m, RSSI fingerprints with an average error of 4.31 *dBm* are generated, and a localization accuracy of 3.9m is achieved based on expanding the fingerprint library.Also to improve the fingerprint localization accuracy, a scheme to build virtual RPs in the localization environment using a signal propagation model is proposed in [33], which is applied to achieve almost the same localization accuracy with 33% less fingerprint point deployment, however, this technique requires additional hardware deployment in the localization area. Reference [30] proposes a fast fingerprint database construction method using adaptive path loss model interpolation, which reduces the workload and construction time by 85%. [34] proposed a virtual fingerprint construction scheme based on a smart antenna system to rapidly construct fingerprint databases with a logarithmic path loss model, reduce the localization error by an area-based remediation algorithm, and achieve an average error of 1.25m in a 6m × 6m localization environment. Reference [35] improved the quality of fingerprints by including the path loss index as part of the construction of fingerprints along with the RSSI values collected or predicted for each RPs.

The above-mentioned methods based on channel propagation models use the actual fingerprint information measured at real points to calculate the environmental loss factor of the wireless signal propagation in a given scenario and estimate the fingerprint information at virtual points. Although such methods can reduce the time and labor cost required for intensive fingerprint collection, their localization accuracy is limited by the degree of prediction and number of signal sources, and the complexity of the localization scenario. With the increase of obstacles (walls, floor slabs) in the localization scene, the signal fading factor calculated by the signal propagation model will no longer be able to comprehensively and accurately characterize the effects of multipath effects, shadow fading, and other factors on wireless signal propagation, and the resulting model is difficult to predict the detailed signal fading and propagation patterns in complex indoor electromagnetic environments. At the same time, the resolution of constructing a channel propagation model to generate virtual point fingerprint information also depends on the number of sources of wireless signals used to construct the signal model, and the number of sources deployed in a building can seriously affect the positioning accuracy. In addition, the signal fading factor and the additional value of wireless signal loss measured and calculated by these methods during the construction of the channel propagation model are scenario-specific and can only achieve effective localization in the scenario where the model is constructed, but cannot be efficiently deployed in other scenarios, and the localization system is not robust. To overcome these problems, researchers have investigated the use of interpolation or regression to generate virtual point fingerprint data, which reduces the dependence on the number of APs and alleviates the degradation of localization accuracy caused by the complexity of scenes.

B. METHODS BASED ON INTERPOLATION OR REGRESSION

As the complexity of indoor scenes increases, there are many indoor objects and complex layout, and the scenes cover a wide area, it is difficult to deploy RPS uniformly in high density to collect fingerprint data and the collection cost is high. Based on interpolation or regression methods, fingerprint points are sparsely deployed and fingerprint data are collected in the location scene, and the distribution information of wireless signals in the unknown area is estimated as fingerprint data based on the distribution of wireless signals at the sparse fingerprint points, so as to quickly build a complete fingerprint database and reduce the cost of intensive fingerprint collection, typical interpolation or regression methods include.

1) LINEAR INTERPOLATION

The linear interpolation method can be considered as the simplest method for predicting the data values of uncollected points. Its principle is based on the estimation of the measured values between two sampled points in the neighborhood of the unsampled points, and the predicted interpolation of the unsampled points is only related to the measured values of the two sampled points, which can be calculated by the following equation (5).

$$Z_0 = w_1 Z_1 + w_2 Z_2 \tag{5}$$

where, Z_0 denotes the prediction interpolation of the uncollected points, Z_1 and Z_2 denote the two sampled point measurements in the neighborhood of Z_0 respectively, and W_1 and W_2 denote the linear interpolation weights.

2) INVERSE DISTANCE WEIGHTED (IDW)

Inverse Distance Weighted (IDW) interpolation is a local deterministic interpolation technique that assigns a value to an unsampled point by a distance-weighted average of the values at the sampled points in the neighborhood of the unsampled point.IDW assumes that sampled points that are closer to the unsampled point have a greater impact on the value of the unsampled point than sampled points that are farther away, so the predicted interpolation is calculated as (6).

$$Z_0 = \frac{\sum_{i=1}^{s} Z_i \frac{1}{d_i^{\theta}}}{\sum_{i=1}^{s} \frac{1}{d_i^{\theta}}} \tag{6}$$

where Z_0 denotes the predicted value of the unsampled point, Z_i denotes the collected value at sample point i, d_i denotes the distance between sample point i and the unsampled point, s denotes the number of sampled points in the neighborhood of the unsampled point, and θ denotes the distance-weighted control parameter.

3) KRIGING INTERPOLATION

Krieger interpolation is similar to IDW in that it also weights the measurements of sampled points around the unsampled points to obtain a predicted interpolation of the unmeasured locations. However, in kriging interpolation, the weights are not only related to the distance between sampled and unsampled points, but are also influenced by the overall spatial distribution of sampled points. The kriging interpolation method quantifies the spatial correlation between the sampled points by a semi-variance function that is fitted to all sampled points in the specified neighborhood of the unsampled points to determine the weight coefficients of each sampled point, and the predicted interpolation is calculated as (7).

$$\hat{Z}(s_0) = \sum_{i=1}^{N} \lambda_i Z(s_i) \tag{7}$$

where $\hat{Z}(s_0)$) denotes the predicted interpolation of the unsampled point S_0 in the specified neighborhood, N denotes the number of sampled points in the specified neighborhood of the unsampled point s_0, $Z(s_i)$ denotes the measured value at the sampled point s_i, and λ_i denotes the weighting factor to be determined.

TABLE 2. Low-cost indoor location method based on interpolation or regression.

Spatial interpolation methods	Ref	Year	Parameter and Technology	Position error(m)	interpolation errors(dBm)	Scene (m^2)
Inverse Distance Weighting	[41]	2012	RSSI, WIFI	×	4.53	5*5
linear interpolation	[38]	2013	RSSI, USRP	5	×	30*30
	[39]	2017	RSSI, WIFI	×	3.9726	85*12
	[40]	2022	RSSI, ZigBee	0.29	×	5*5
kriging interpolation	[42]	2015	RSSI, WIFI	<5	×	150*150
	[43]	2016	RSSI, WIFI	1.265	2	27*8
	[22]	2018	RSSI, iBeacon	<3	<3	31*22
	[44]	2020	RSSI, WIFI	2.86	×	120*245
cubic spline interpolation	[28]	2018	RSSI, WIFI	<2	×	49.4*14.1
Regression	[46]	2013	RSSI, WIFI	<6	×	70*50
	[48]	2020	RSSI, iBeacon	2.17	×	38*14
	[47]	2021	RSSI, WIFI	3.94	3.66	50*36

× : not covered in the article

4) GAUSSIAN PROCESS REGRESSION

Gaussian process regression is a Bayesian nonparametric regression method that predicts the observations at unsampled points $y(z_0)$ from the posterior probability distribution over the mapping function inferred from the measurements at the sampled points. The process first assumes a Gaussian prior distribution of the mapping function parameters and updates the probabilities using Bayes' theorem to model the set of observations at N sampled points as a multivariate Gaussian distribution, where the prior distribution is defined by the mean function $\mu(\cdot)$) and the covariance or kernel function $K(\cdot,\cdot)$), so the Gaussian regression process can be expressed as (8)

$$y(z_0) \sim \mathrm{GPR}\left[\mu(\cdot), K(\cdot,\cdot;\boldsymbol{\theta}) \mid z_0\right] \tag{8}$$

where z_0 denotes the unsampled points to be interpolated and θ is the hyperparameter of the kernel function, which is usually trained by maximizing the marginal log probability using the training dataset.

As shown in Table 2, various methods of interpolation or regression, such as linear interpolation, cubic spline interpolation, kriging interpolation, inverse distance weighted interpolation, and Gaussian process regression, have been applied to generate the complete fingerprint database.

Interpolation or regression is the process of estimating missing values from a set of known values. Reference [38] improved the accuracy, correctness, and robustness of fingerprint localization systems in complex environments by using linear interpolation to calculate fingerprint information for a specific fingerprint point based on a set of neighboring fingerprint points using spatial correlation of neighboring locations. Reference [39] studied the use of linear interpolation for missing fingerprints to reduce the time required to create a fingerprint database, and compared the RSSI generated values with the actual measured values with an average error of up to 3.97dbm. Reference [40] used a relatively sparse RSSI fingerprint dataset, applied a bilinear interpolation technique to construct a generative database, and built a fingerprint localization system based on the generative database with an average localization error of only 0.29m in a 5m $\times$ 5m localization environment.The basic idea of IDW is to provide weights for data points based on their distances to the estimation points. It is known that the closer the data point is to the estimation point, the larger its weight is [31]. Reference [41] used an adaptive smoothing algorithm to deal with RSSI discontinuities due to the presence of interior walls, which in turn generated a complete fingerprint database using the IDW method. Kriging interpolation is a geostatistical method used to make optimal spatial predictions at unobserved locations and has some advantages over some other spatial interpolation methods. References [42] and [43] used kriging interpolation to construct a complete interpolation database for target localization based on sparsely collected fingerprint data. Reference [22] used kriging interpolation to generate RSSI fingerprint data in the region of undeployable fingerprint collection points to cover and enhance the entire fingerprint database, achieving a localization accuracy of less than 3m in a 31m $\times$ 22m localization environment. Reference [44] proposed a proximity relationship-based indoor localization method under sparse fingerprint data, using kriging interpolation to enrich sparse fingerprint data, setting proximity relationship boundaries to restrict fingerprint data generation to reduce fingerprint ambiguity and improve fingerprint matching efficiency, achieving a localization accuracy of 2.86 m at an interpolation density of 1 m. Reference [46] proposed a method to construct a generative database based on Gaussian process regression, using both real and virtual data, to improve the accuracy of the localization system. Reference [47] applied regression techniques to analyze RSSI signal features and generate virtual data to solve the richness problem of fingerprint database. Reference [48] proposed a fast fingerprint construction method based on Gaussian process regression (GPR) to sparsely sample RSSI in space to generate a full-space fingerprint database, while reducing outliers to improve fingerprint database reliability, and achieved a localization accuracy of 2.17m average error in a 38m $\times$ 14m localization environment.

The above-mentioned interpolation or regression-based methods use fingerprint data collected at real points to generate fingerprint information at virtual points, expanding the density of fingerprint points in the localization environment and increasing the richness of the fingerprint database, which in turn increases the accuracy of fingerprint localization methods, while such methods can be rapidly deployed with better robustness by only collecting fingerprint data at real points in a new scene. However, there are still limitations in these methods, as the interpolation or regression methods only use local information to infer the wireless signal information at the virtual point, and do not consider the effect of global information. Moreover, the localization accuracy of these methods is also affected by the deployment of real points. Different distribution patterns of real points will generate different virtual fingerprint information and obtain different interpolation results, so whether the real points can be deployed and information collected according to the optimal spatial pattern will seriously affect the effectiveness of these methods.

The methods of generating virtual fingerprints by traditional data enhancement means are all based on the fingerprints of partial collection points to generate fingerprint information of uncollected areas, and reduce the labor and time costs by reducing the number of fingerprint collection points. However, these methods can only reduce the number of fingerprint points and generate virtual fingerprint points, but cannot increase the number of samples for each fingerprint collection point. The fingerprint database constructed by these methods is difficult to comprehensively cover the fingerprint features under multiple spatial and temporal conditions, and will differ from the online data distribution when the localization environment changes dynamically, leading to a decrease in localization accuracy. And for each real point, the computation of collecting a large number of fingerprints to support model construction, interpolation or regression still takes a lot of time.

III. ADVANCED APPROACHES

Wireless signal strength (RSSI) or wireless channel state information (CSI) is mostly used as fingerprint data in fingerprint localization, however, as coarse-grained physical layer information, wireless signal propagation is affected by various factors such as multipath fading, temperature and humidity variations, door opening and closing, furniture repositioning, and the presence and movement of people [49]. Due to these factors, RSSI and CSI are still highly variable at fixed locations. Therefore, fingerprint localization systems should have the ability to adapt to dynamic changes in the localization scene and resist fluctuations in the online fingerprint distribution caused by various factors. To offset the channel variability caused by several propagation factors such as fading, shadowing, and multipath effects, and to ensure that the fingerprint database covers as much as possible the distribution of fingerprint information in the location scene under multiple spatial and temporal conditions, fingerprint data from each fingerprint point needs to be repeatedly measured several times when the environment and infrastructure of the location scene and the density of people change. In addition, when the positioning scene is changed, the changes in space size, signal source location and number, and environmental

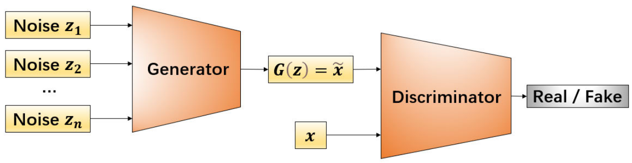


FIGURE 2. Generate adversarial network models.

layout will cause the fingerprint database of the old scene to fail to match the fingerprint distribution of the new scene, requiring a large amount of fingerprint data to be re-collected in the new scene [50], which will consume a large amount of labor and time costs, prolong the deployment cycle of the fingerprint positioning system, and reduce the feasibility of widespread deployment.

Advanced machine learning models have been used to overcome these challenges by combining indoor localization with artificial intelligence. Generative adversarial networks, which are widely used for data augmentation, can both learn the features of existing fingerprint data to expand the size of the data base by generating virtual fingerprint point data and generate fingerprint data from existing fingerprint points to increase the diversity of fingerprint samples, significantly reducing the labor cost required to collect fingerprints. The few-shot learning network, which achieves better results in few-shot classification tasks, achieves high-precision localization results in new scenes with only a small amount of collected data by reusing the complete data set of existing localization scenes, reducing the time cost of fingerprint collection required to redeploy the fingerprint localization system for new scenes. In this section, two advanced machine learning methods based on 1) generative adversarial networks and 2) few-shot learning are presented to achieve efficient localization methods using fingerprint data.

A. METHODS BASED ON GENERATIVE ADVERSARIAL NETWORKS

Generative adversarial networks (GANs) have shown good capabilities in generating various types of data as additional measurements [51], [52], such as images, text, sound, etc., significantly reducing acquisition time and saving labor costs.

Generative adversarial network (GAN), inspired by game theory, uses an adversarial dual network structure to estimate the distribution probabilities of real data sets during training and is a generative model consisting of a generator G and a discriminator D. The underlying architecture is shown in Figure 2 [53]. Where the generator G generates fake data as close as possible to the probability of the distribution of the real data during the training process, while the discriminator D correctly distinguishes between real and fake data as much as possible during the training process, so that when one of the networks is optimized, the other one is improved as well. The input of generator G is usually a random noise vector z with uniform or normal distribution, and the network through generator G maps the noise z to a new data space to obtain fake data $G(Z)$ with the same dimension as the real data. The discriminator D is a binary classifier that takes the real data in the dataset and the fake data generated by the generator G as input, and its output indicates whether the input samples are real data or not. When the discriminator D cannot determine whether the input samples are from the real dataset or the generator G, the generative adversarial network reaches its optimal state, when the generator model G can already generate pseudo-data with the same distribution as the real data.

As two important network structures in generative adversarial networks, generator G and discriminator D have their own loss functions $J^{(G)}$ and $J^{(D)}$, respectively. As a binary classifier, the discriminator D has the following loss function as (9).

$$J^{(D)} = -\frac{1}{2}\mathbb{E}_{x\sim p}\log D(x) - \frac{1}{2}\mathbb{E}_{z}\log(1-D(G(z))) \tag{9}$$

where x denotes the sample from the real data set, z denotes the random noise vector input to generator G, $G(z)$ denotes the fake data generated by generator G, $\mathbb{E}$ denotes the expected value, $D(x)$ denotes the probability that discriminator D will identify x as real data, and $D(G(z))$ denotes the probability that discriminator D will identify the data generated by generator G as real data. The training goal of the discriminator is to correctly distinguish whether the input samples are true data, so $J^{(D)}$ keeps $D(G(z))$ close to 0 during the training process, while the goal of the generator G is to make its generated data can be identified as true data by the discriminator D, so $J^{(G)}$ keeps $D(G(z))$ close to 1 during the training process, so the loss function of the generator is (10)

$$J^{(G)} = -J^{(D)} \tag{10}$$

Therefore, the loss function of generating adversarial network can be transformed by combining the conflict between $J^{(G)}$ and $J^{(D)}$ into (11).

$$\min_{G}\max_{D} V(D,G) = \mathbb{E}_{x\sim p(x)}[\log D(x)] + \mathbb{E}_{z\sim p(z)}[\log(1-D(G(z)))] \tag{11}$$

During the training process, the network parameters in the generator G are updated together with the network parameters in the discriminator D. When the discriminator D cannot determine whether the data generated by the generator G is false data, namely $D(G(z)) = 0.5$, the generative adversarial network will be able to generate fake data with approximately the same distribution as the real sample, effectively realizing data expansion.

Therefore, depending on the fingerprint data collection method used for training, the use of GAN can generate both pseudo-fingerprint data with the same fingerprint points as the training data to increase the diversity of fingerprint samples and virtual fingerprint data at non-training fingerprint points to expand the fingerprint database size [54]. The use of fingerprint data belonging to multiple fingerprint points

TABLE 3. Low-cost indoor location method based on Generative Adversarial network.

Ref	Year	Parameter and Technology	Position error (m)	Scene (m^2)
[55]	2019	CSI, WFIF	0.92	7*7
[56]	2020	CSI,WIFI	×	8.6*5.8
[60]	2020	RSSI, WIFI	1.98	20*20
[57]	2021	CSI, WIFI	0.99	12.6*4.8
[58]	2021	RSSI, WIFI	3.91	×
[59]	2021	RSSI, WIFI	0.83	20*20

× : It is not covered in the article

with diversity to build a complete dataset can support the fingerprint localization system to maintain good performance in scenarios with complex layouts and frequent changes. As shown in Table 3, the use of GAN to generate virtual fingerprint data to increase the richness of fingerprint database and improve the localization accuracy has gradually become a research hotspot in the field of indoor localization in recent years.

Reference [55] extended the initial training fingerprint database based on DCGAN by increasing the amount of training data collected at each reference point to enhance the fingerprint database richness, and achieved a localization accuracy of 0.92m average error in a 7m × 7m localization environment. Reference [56] used GAN for semi-supervised learning, where GAN used labeled and unlabeled data to share weights with the localization classifier, benefiting from unlabeled data when labeled data was insufficient. Reference [57] constructed CSI fingerprint data at each location point as amplitude images and then generated additional images to extend the fingerprint dataset by AC-GAN, which can generate pseudo-fingerprint data without limitation and efficiently utilizes computational resources and real fingerprint data. Reference [58] generated pseudo-fingerprint data at different locations based on CGAN using real collected RSSI data. Reference [59] proposed a generative adversarial network for RSSI data enhancement, which generates virtual RSSI data based on a small amount of collected marker data and selects the generated data to reduce data generation errors, with an average localization error of up to 0.83m in a 20m × 20m localization environment.

The aforementioned generative adversarial network-based fingerprint localization method generates fingerprint data for virtual fingerprint points, which achieves the effect of expanding the fingerprint database, while being able to generate fingerprint data for multiple spatio-temporal conditions at fingerprint points, increasing the diversity of fingerprint samples and making the fingerprint database contain a rich distribution of fingerprint information. The method efficiently utilizes existing fingerprint data to overcome the impact of wireless signal variability on fingerprint localization accuracy, improves the robustness of the localization system to environmental dynamics [59], significantly reduces the labor and time costs required to collect fingerprint data repeatedly in large quantities, and increases the feasibility of deploying indoor localization systems in complex indoor environments. However, the generative adversarial network-based fingerprint localization method is only capable of generating fingerprint data in a single localization environment without multi-environment robustness, and when the localization environment changes, it is still necessary to re-collect fingerprint data and retrain the generative network in the new environment, so the deployment in the new environment still suffers from high labor costs.

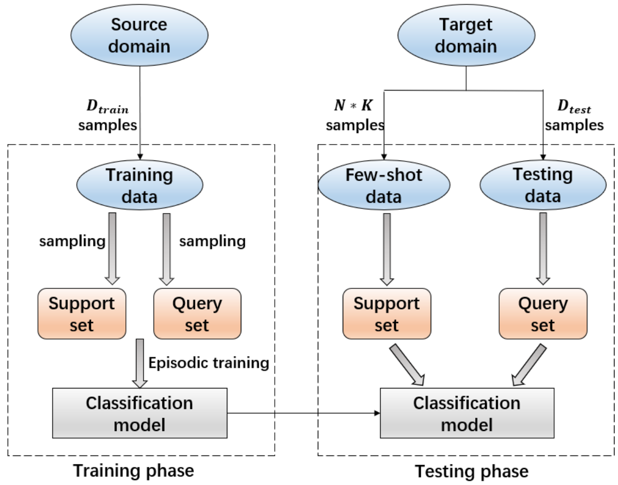


FIGURE 3. The general framework of few-shot learning.

B. METHODS BASED ON FEW-SHOT LEARNING

Few-shot learning, a popular field of machine learning in recent years, has been favored by researchers for its advantages such as no large-scale training data and small learning cost [61], [62]. Applying few-shot learning to solve classification tasks can reduce the cost of large-scale collection and labeling of data, and thus has a very promising application in areas where data are scarce and difficult to collect [63]. few-shot learning is dedicated to training network models in datasets with a large amount of labeled data already available (source domain) and categorizing unlabeled data samples in datasets with only a small amount of labeled data (target domain) [64], [65], and its general framework is shown in Figure 3.

Few-shot learning trains a network model in the training phase so that it can achieve the task effect well when the target task has only a few labeled samples. In the training stage, given a training dataset D_{train}, we use a process called episodic to sample the support set S and query set Q from the episodic data set. That is, the samples of support set S and query set Q have data labels.Where the support set $S \subset D_{\text{train}}$ contains N classes and each class has K samples (namely N-way K-shot), which can be represented as $S = \{(x_1, y_1), (x_2, y_2), \ldots, (x_{N\times K}, y_{N\times K})\}$.Query set $Q \subset D_{\text{train}}$ contains $\bar{D}$ samples, which can be denoted as $Q = \left\{(x_{N\times K+1}, y_{N\times K+1}), \ldots, \left(x_{N\times K+\bar{D}}, y_{N\times K+\bar{D}}\right)\right\}$. D_{test}

in the testing phase, a given test data sets, with training set of few-shot learning model can support sample a few cases, accurately map the query set $Q \subset D_{\text{test}}$ no labels of the samples to support on the similar sample similar tags, it is worth noting that the training data set and test data set of tags are mutually exclusive, the $D_{\text{train}} \cap D_{\text{test}} = \emptyset$.

Few-shot learning has been used with good results in classification tasks in several fields. References [66] and [67] designed few-shot learning models based on graph convolutional neural networks to achieve good results in image classification using only 1, 5, and 10 samples per class of images. Reference [68] designed a classification algorithm based on few-shot learning to significantly improve the performance of gesture classifier using a weakly supervised training dataset containing multiple gesture instances, and achieved good classification results on a large-scale gesture recognition dataset using only a single gesture sample for training. Reference [69] proposed a composite memory network (CMN) structure based on the key-value memory network paradigm for few-shot video classification, and demonstrated the effectiveness of the method on a few-shot video classification dataset. Reference [70] to solve the drug discovery problem, LSTM was introduced into a few-shot learning network based on graph neural networks, which improved the ability to exchange feature information among drug small molecules during training and significantly reduced the amount of data required to perform drug efficacy prediction in drug discovery experiments. Reference [71] proposed an adaptive metric-based few-shot learning method that can obtain a set of metric values during the training phase and automatically determine the best combination of weights to complete the task in the testing phase. Significant prediction results were achieved on sentiment analysis and conversational intent classification datasets.

Therefore, it is feasible to use a few-shot learning model to reuse the existing complete data set in the old localization environment and use only a small amount of fingerprint data collected in the new environment for localization, reducing the cost of collecting and tagging fingerprint data for localization in the new localization scenario. In [50], the concept of few-shot learning is first introduced to indoor localization, and the ''task'' of few-shot learning is specified as localization in various heterogeneous environments. When the settings, such as layout, size, and number of fingerprint points deployed, are different in the old and new localization environments, the method can show significant localization results with only 1, 5, or 10 marked CSI fingerprint samples collected in the offline phase for each fingerprint point in the new environment. The performance is consistent with that achieved by a convolutional neural network-based localization model trained to collect 400 labeled CSI fingerprint samples per fingerprint point in the new environment, with a more than 40-fold difference in fingerprint data collection requirements between the two. The method significantly reduces the labor and time costs of fingerprint data acquisition and labeling required for localization in multiple different indoor environments, increasing the feasibility of rapid deployment of fingerprint localization systems in multiple localization environments.

In conclusion, fingerprint localization methods based on few-shot learning can efficiently utilize fingerprint data, effectively reduce the labor and time costs required for repeated fingerprint data acquisition in localization scenarios, and increase the robustness of fingerprint localization systems. Although studies have been conducted to combine few-shot learning with the field of indoor localization to reduce the labor consumption required to repeatedly collect large amounts of fingerprint data in new localization environments, this direction is still among the new ideas to overcome indoor fingerprint localization challenges and has full potential to improve the feasibility of rapid multi-scene deployment of indoor fingerprint localization systems. For example, a) how to extract stable localization fingerprint features from a small number of support set samples to cope with the degradation of localization performance due to fluctuations in fingerprint distribution as the environment dynamically changes in a new scene. b) how to distinguish similar fingerprints collected from two different locations based on a small number of sample data in the support set due to multipath reflections of wireless signals in a complex indoor environment with blurred fingerprint data [72]. Solving the above problems is the direction and challenge for further research on fingerprint localization based on few-shot learning. With the continuous combination of machine learning and indoor localization field, it is foreseen that fingerprint localization technology based on few-shot learning is one of the development trends of indoor localization technology in the future.

IV. CONCLUSION

In this review, we provide a comprehensive overview of indoor localization methods for building wireless fingerprint databases at low cost. Representative methods proposed in recent years are comprehensively reviewed, compared and summarized according to classical data augmentation methods and advanced machine learning models, respectively. For each class of methods, we elaborate their underlying models or algorithms as theoretical introduction and practical guidelines. Finally, we outline current challenges and future research directions, and show that there is much room for further exploration. Overall, we hope that this review can serve as a guide for indoor positioning researchers to address the high cost of data acquisition and to advance research and development in this area.

REFERENCES

[1] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2568–2599, 3rd Quart., 2017.

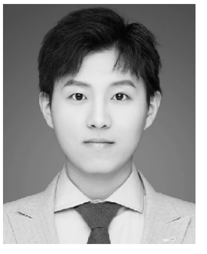
[2] A. Yassin, "Recent advances in indoor localization: A survey on theoretical approaches and applications," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 1327–1346, 2nd Quart., 2016.

[3] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Trans. Syst., Man Cybern., C, Appl. Rev.*, vol. 37, no. 6, pp. 1067–1080, Nov. 2007.

[4] V. Moghtadaiee and A. G. Dempstera, "Indoor location fingerprinting using FM radio signals," *IEEE Trans. Broadcast.*, vol. 60, no. 2, pp. 336–346, Jun. 2014.
[5] 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 *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Sep. 2019, pp. 1–8.
[6] J. Bi, M. Zhao, G. Yao, H. Cao, Y. Feng, H. Jiang, and D. Chai, "PSOSVR-Pos: WiFi indoor positioning using SVR optimized by PSO," *Expert Syst. Appl.*, vol. 222, Jul. 2023, Art. no. 119778.
[7] A. Khalajmehrabadi, N. Gatsis, and D. Akopian, "Modern WLAN fingerprinting indoor positioning methods and deployment challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1974–2002, 3rd Quart., 2017.
[8] F. Zampella, A. R. J. Ruiz, and F. S. Granja, "Indoor positioning using efficient map matching, RSS measurements, and an improved motion model," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1304–1317, Apr. 2015.
[9] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Comput. Surv.*, vol. 46, no. 2, pp. 1–32, Nov. 2013.
[10] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, and L. M. Ni, "CSI-based Indoor Localization," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 7, pp. 1300–1309, Jul. 2013.
[11] M. Kessel and M. Werner, "SMARTPOS: Accurate and precise indoor positioning on mobile phones," in *Proc. 1st Int. Conf. Mobile Services, Resour., Users (MOBILITY)*, 2011.
[12] S. Tsruya, "Devices, methods, and systems for radio map generation," U.S. Patent 8 938 255, Jan. 20, 2015.
[13] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM Conf. Comput. Commun. 19th Annu. Joint Conf. IEEE Comput. Commun. Societies*, Mar. 2000, pp. 775–784.
[14] F. Parralejo, F. J. Aranda, J. A. Paredes, F. J. Alvarez, and J. Morera, "Comparative study of different BLE fingerprint reconstruction techniques," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Nov. 2021, pp. 1–8.
[15] J. Riihijarvi, P. Mahonen, M. Wellens, and M. Gordziel, "Characterization and modelling of spectrum for dynamic spectrum access with spatial statistics and random fields," in *Proc. IEEE 19th Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2008, pp. 1–6.
[16] C. Li, Q. Xu, Z. Gong, and R. Zheng, "TuRF: Fast data collection for fingerprint-based indoor localization," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Sep. 2017, pp. 1–8.
[17] J. Jun, L. He, Y. Gu, W. Jiang, G. Kushwaha, A. Vipin, L. Cheng, C. Liu, and T. Zhu, "Low-overhead WiFi fingerprinting," *IEEE Trans. Mobile Comput.*, vol. 17, no. 3, pp. 590–603, Mar. 2018.
[18] W. Sun, X. Yuan, J. Wang, Q. Li, L. Chen, and D. Mu, "End-to-end data delivery reliability model for estimating and optimizing the link quality of industrial WSNs," *IEEE Trans. Autom. Sci. Eng.*, vol. 15, no. 3, pp. 1127–1137, Jul. 2017.
[19] J. Hu, "Experimental analysis on weight K-nearest neighbor indoor fingerprint positioning," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 891–897, Feb. 2018.
[20] Z. Yang, C. Wu, and Y. Liu, "Locating in fingerprint space: Wireless indoor localization with little human intervention," in *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw.*, Aug. 2012, pp. 269–280.
[21] A. Nessa, B. Adhikari, F. Hussain, and X. N. Fernando, "A survey of machine learning for indoor positioning," *IEEE Access*, vol. 8, pp. 214945–214965, 2020.
[22] J. Zuo, S. Liu, H. Xia, and Y. Qiao, "Multi-phase fingerprint map based on interpolation for indoor localization using iBeacons," *IEEE Sensors J.*, vol. 18, no. 8, pp. 3351–3359, Apr. 2018.
[23] X. Tong, Y. Wan, Q. Li, X. Tian, and X. Wang, "CSI fingerprinting localization with low human efforts," *IEEE/ACM Trans. Netw.*, vol. 29, no. 1, pp. 372–385, Feb. 2021.
[24] B. Sklar, "Rayleigh fading channels in mobile digital communication systems. I. Characterization," *IEEE Commun. Mag.*, vol. 35, no. 7, pp. 90–100, Jul. 1997.
[25] V. Honkavirta, T. Perala, S. Ali-Loytty, and R. Piche, "A comparative survey of WLAN location fingerprinting methods," in *Proc. 6th Workshop Positioning, Navigat. Commun.*, Mar. 2009, pp. 243–251.
[26] B. Li, Y. Wang, H. K. Lee, A. Dempster, and C. Rizos, "Method for yielding a database of location fingerprints in WLAN," *IEE Proc.-Commun.*, vol. 152, no. 5, pp. 580–586, Oct. 2005.
[27] M. M. Atia, A. Noureldin, and M. J. Korenberg, "Dynamic online-calibrated radio maps for indoor positioning in wireless local area networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 9, pp. 1774–1787, Sep. 2013.
[28] L. Ma, W. Zhao, Y. Xu, and C. Li, "Radio map efficient building method using tensor completion for WLAN indoor positioning system," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.
[29] R. Kubota, S. Tagashira, Y. Arakawa, T. Kitasuka, and A. Fukuda, "Efficient survey database construction using location fingerprinting interpolation," in *Proc. IEEE 27th Int. Conf. Adv. Inf. Netw. Appl. (AINA)*, Mar. 2013, pp. 469–476.
[30] J. Bi, Y. Wang, Z. Li, S. Xu, J. Zhou, M. Sun, and M. Si, "Fast radio map construction by using adaptive path loss model interpolation in large-scale building," *Sensors*, vol. 19, no. 3, p. 712, Feb. 2019.
[31] V. Moghtadaiee, S. A. Ghorashi, and M. Ghavami, "New reconstructed database for cost reduction in indoor fingerprinting localization," *IEEE Access*, vol. 7, pp. 104462–104477, 2019.
[32] T. Guan, L. Fang, W. Dong, D. Koutsonikolas, G. Challen, and C. Qiao, "Robust, cost-effective and scalable localization in large indoor areas," *Comput. Netw.*, vol. 120, pp. 43–55, Jun. 2017.
[33] S.-T. Sheu, Y.-M. Hsu, and H.-Y. Chen, "Indoor location estimation using smart antenna system with virtual fingerprint construction scheme," in *Proc. Int. Conf. Mobile Ubiquitous Comput., Syst., Services Technol. (UBICOMM)* 2014, pp. 1–6.
[34] Y.-H. Wu, Y.-L. Chen, and S.-T. Sheu, "Indoor location estimation using virtual fingerprint construction and zone-based remedy algorithm," in *Proc. Int. Conf. Commun. Problem-Solving (ICCP)*, Sep. 2016, pp. 1–3.
[35] J. Zhang, G. Han, N. Sun, and L. Shu, "Path-loss-based fingerprint localization approach for location-based services in indoor environments," *IEEE Access*, vol. 5, pp. 13756–13769, 2017.
[36] L. F. M. de Moraes and B. A. A. Nunes, "Calibration-free WLAN location system based on dynamic mapping of signal strength," in *Proc. 4th ACM Int. Workshop Mobility Manage. Wireless Access*, Oct. 2006, pp. 92–99.
[37] K. Chintalapudi, A. P. Iyer, and V. N. Padmanabhan, "Indoor localization without the pain," in *Proc. 16th Annu. Int. Conf. Mobile Comput. Netw.*, Sep. 2010, pp. 173–184.
[38] C. Koweerawong, K. Wipusitwarakun, and K. Kaemarungsi, "Indoor localization improvement via adaptive RSS fingerprinting database," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, Jan. 2013, pp. 412–416.
[39] J. Racko, J. Machaj, and P. Brida, "Wi-Fi fingerprint radio map creation by using interpolation," *Proc. Eng.*, vol. 192, pp. 753–758, Jan. 2017.
[40] D. J. Suroso, F. Y. M. Adiyatma, P. Cherntanomwong, and P. Sooraksa, "Fingerprint database enhancement by applying interpolation and regression techniques for IoT-based indoor localization," *Emerg. Sci. J.*, vol. 4, pp. 167–189, Jan. 2022.
[41] W. Bong and Y. C. Kim, "Fingerprint Wi-Fi radio map interpolated by discontinuity preserving smoothing," in *Proc. Int. Conf. Hybrid Inf. Technol.* Berlin, Germany: Springer, 2012, pp. 138–145.
[42] C. Liu, A. Kiring, N. Salman, L. Mihaylova, and I. Esnaola, "A Kriging algorithm for location fingerprinting based on received signal strength," in *Proc. Sensor Data Fusion: Trends, Solutions, Appl. (SDF)*, Oct. 2015, pp. 1–6.
[43] H. Zhao, B. Huang, and B. Jia, "Applying Kriging interpolation for WiFi fingerprinting based indoor positioning systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2016, pp. 1–6.
[44] Y. Wang, R. Guo, W. Wang, X. Li, S. Tang, W. Zhang, L. Wang, L. Chen, Y. Li, and W. Xiu, "Near relation-based indoor positioning method under sparse Wi-Fi fingerprints," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 12, p. 714, Dec. 2020.
[45] S. Han, Y. Li, W. Meng, C. Li, T. Liu, and Y. Zhang, "Indoor localization with a single Wi-Fi access point based on OFDM-MIMO," *IEEE Syst. J.*, vol. 13, no. 1, pp. 964–972, Mar. 2019.
[46] Y. Cho, J. Kim, M. Ji, Y. Lee, and S. Park, "GPR based Wi-Fi radio map construction from real/virtual indoor dynamic surveying data," in *Proc. 13th Int. Conf. Control, Autom. Syst. (ICCAS)*, Oct. 2013, pp. 712–714.
[47] G. M. Mendoza-Silva, A. C. Costa, J. Torres-Sospedra, M. Painho, and J. Huerta, "Environment-aware regression for indoor localization based on WiFi fingerprinting," *IEEE Sensors J.*, vol. 22, no. 6, pp. 4978–4988, Mar. 2022.
[48] H. Ai, K. Tang, W. Huang, S. Zhang, and T. Li, "Fast fingerprints construction via GPR of high spatial-temporal resolution with sparse RSS sampling in indoor localization," *Computing*, vol. 102, no. 3, pp. 781–794, Mar. 2020.

[49] H. Lim et al., ''Zero-configuration, robust indoor localization: Theory and experimentation,'' in *Proc. 25th IEEE Int. Conf. Comput. Commun. (INFOCOM)*, 2007.
[50] B.-J. Chen and R. Y. Chang, ''Few-shot transfer learning for device-free fingerprinting indoor localization,'' 2022, *arXiv:2201.12656*.
[51] I. Goodfellow, ''Generative adversarial networks,'' *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.
[52] Y.-J. Cao, L.-L. Jia, Y.-X. Chen, N. Lin, C. Yang, B. Zhang, Z. Liu, X.-X. Li, and H.-H. Dai, ''Recent advances of generative adversarial networks in computer vision,'' *IEEE Access*, vol. 7, pp. 14985–15006, 2019.
[53] Z. Pan, W. Yu, X. Yi, A. Khan, F. Yuan, and Y. Zheng, ''Recent progress on generative adversarial networks (GANs): A survey,'' *IEEE Access*, vol. 7, pp. 36322–36333, 2019.
[54] A. Belmonte-Hernandez, G. Hernandez-Penaloza, D. M. Gutierrez, and F. Alvarez, ''Recurrent model for wireless indoor tracking and positioning recovering using generative networks,'' *IEEE Sensors J.*, vol. 20, no. 6, pp. 3356–3365, Mar. 2020.
[55] Q. Li, ''AF-DCGAN: Amplitude feature deep convolutional GAN for fingerprint construction in indoor localization systems,'' *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 5, no. 3, pp. 468–480, Jun. 2019.
[56] K. M. Chen and R. Y. Chang, ''Semi-supervised learning with GANs for device-free fingerprinting indoor localization,'' in *Proc. GLOBECOM IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.
[57] W. Wei, J. Yan, L. Wan, C. Wang, G. Zhang, and X. Wu, ''Enriching indoor localization fingerprint using a single AC-GAN,'' in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2021, pp. 1–6.
[58] W. Njima, M. Chafii, and R. M. Shubair, ''GAN based data augmentation for indoor localization using labeled and unlabeled data,'' in *Proc. Int. Balkan Conf. Commun. Netw. (BalkanCom)*, Sep. 2021, pp. 36–39.
[59] W. Njima, M. Chafii, A. Chorti, R. M. Shubair, and H. V. Poor, ''Indoor localization using data augmentation via selective generative adversarial networks,'' *IEEE Access*, vol. 9, pp. 98337–98347, 2021.
[60] H. Zou, ''Adversarial learning-enabled automatic WiFi indoor radio map construction and adaptation with mobile robot,'' *IEEE Internet Things J.*, vol. 7, no. 8, pp. 6946–6954, Aug. 2020.
[61] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, ''Generalizing from a few examples: A survey on few-shot learning,'' *ACM Comput. Surveys*, vol. 53, no. 3, pp. 1–34, May 2021.
[62] Y. Wang et al., ''Generalizing from a few examples: A survey on few-shot learning,'' *ACM Comput. Surv.*, vol. 53, no. 3, pp. 1–34, 2020.
[63] A. Parnami and M. Lee, ''Learning from few examples: A summary of approaches to few-shot learning,'' 2022, *arXiv:2203.04291*.
[64] O. Vinyals, ''Matching networks for one shot learning,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, 2016, pp. 1–9.
[65] Q. Sun, Y. Liu, T.-S. Chua, and B. Schiele, ''Meta-transfer learning for few-shot learning,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 403–412.
[66] J. Kim, T. Kim, S. Kim, and C. D. Yoo, ''Edge-labeling graph neural network for few-shot learning,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 11–20.
[67] L. Yang, L. Li, Z. Zhang, X. Zhou, E. Zhou, and Y. Liu, ''DPGN: Distribution propagation graph network for few-shot learning,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 13390–13399.
[68] T. Pfister, J. Charles, and A. Zisserman, ''Domain-adaptive discriminative one-shot learning of gestures,'' in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2014, pp. 814–829.
[69] L. Zhu and Y. Yang, ''Compound memory networks for few-shot video classification,'' in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 751–766.
[70] H. Altae-Tran, B. Ramsundar, A. S. Pappu, and V. Pande, ''Low data drug discovery with one-shot learning,'' *ACS Central Sci.*, vol. 3, no. 4, pp. 283–293, 2017.
[71] M. Yu, X. Guo, J. Yi, S. Chang, S. Potdar, Y. Cheng, G. Tesauro, H. Wang, and B. Zhou, ''Diverse few-shot text classification with multiple metrics,'' 2018, *arXiv:1805.07513*.
[72] B. Huang, Z. Xu, B. Jia, and G. Mao, ''An online radio map update scheme for WiFi fingerprint-based localization,'' *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6909–6918, Apr. 2019.

WEN LIU was born in 1967. She received the B.S. degree from the Xi'an University of Technology, in 1990, and the Ph.D. degree from the Beijing University of Posts and Telecommunications, in 2013. She is currently a Senior Engineer with the Wireless Network Positioning and Communication Fusion Research Center, School of Electronics Engineering, Beijing University of Posts and Telecommunications. Her research interests include indoor positioning, wireless sensor networks, and multi-integrated positioning.



YINGENG ZHANG was born in Beijing, China, in 1999. He received the B.S. degree in electronic information engineering from the Beijing University of Posts and Telecommunications, Beijing, in 2021, where he is currently pursuing the M.S. degree in electronic information engineering. His current research interest includes indoor fingerprint location.

ZHONGLIANG DENG (Senior Member, IEEE) was born in 1965. He received the M.S. degree from the Beijing University of Aeronautics and Astronautics, in 1991, and the Ph.D. degree from Tsinghua University, in 1994. He is currently a Professor with the Wireless Network Positioning and Communication Fusion Research Center, School of Electronics Engineering, Beijing University of Posts and Telecommunications. His research interests include satellite positioning, indoor positioning, and micro-nano systems.

HEYANG ZHOU received the B.S. degree in electronic science and technology from the Beijing University of Posts and Telecommunications, Beijing, China, in 2020, where he is currently pursuing the M.S. degree with the Department of Electrical Engineering. His research interests include fingerprint positioning and multipath suppression algorithm.