IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 25 July 2023, accepted 8 August 2023, date of publication 11 August 2023, date of current version 16 August 2023.
Digital Object Identifier 10.1109/ACCESS.2023.3304334

RESEARCH ARTICLE

The Improvement and Comparison Study of Distance Metrics for Machine Learning Algorithms for Indoor Wi-Fi Localization

XINYUE WANG
Department of Electrical Engineering and Information Technology, Shandong University of Science and Technology, Jinan 250031, China
e-mail: 201903203122@sdust.edu.cn

ABSTRACT Accurate indoor positioning is crucial for many location-based services, but GPS accuracy is significantly reduced due to issues such as signal penetration and accuracy in indoor scenarios. In contrast, indoor Wi-Fi positioning is emerging as a promising alternative in the field. This study proposes a model that combines the k-nearest neighbor algorithm with the dynamic time regularization distance metric for indoor Wi-Fi positioning, and investigates methods for optimizing this model. The traditional K-nearest neighbor algorithm usually uses Euclidean distance for distance calculation, which has the disadvantage of being affected by the length of the signal sequence, resulting in inaccurate calculation of the distance between adjacent points with different time intervals. The dynamic time regularization is more suitable for signals of different lengths like Wi-Fi, which can bend the time axis to make the alignment of two Wi-Fi sequences more accurate. Using DTW as the distance measure of KNN is DTW-KNN. In addition, to enhance the model's ability to handle large-scale data sets, We use Gaussian sum matrices instead of the distance matrix of the traditional dynamic time regularization algorithm. Once again, the standard deviation sigma of the Gaussian distribution and the distance hyperparameters of the K-nearest neighbors are optimally chosen for the most suitable values of Wi-Fi signals. Finally, a fast recognition model based on intermittent downsampling and an accurate recognition model with complete sampling are designed to cope with the focus on real-time and accuracy in different scenarios. These two models can achieve 95.3% and 98% accuracy, respectively, on the public dataset (Wireless Indoor Localization) of indoor Wi-Fi localization.

INDEX TERMS Indoor positioning, Wi-Fi signal processing, machine learning, model optimization, dynamic time warping, Gaussian kernel matrix.

I. INTRODUCTION

Indoor positioning problem has been a persistent and challenging issue, which has yet to be fully resolved despite considerable advancements in outdoor position aiding [1]. With the growing interest in developing location-based services in indoor environments [2], the importance of addressing this issue has become increasingly apparent. While the principles and technologies underlying indoor and outdoor positioning systems are analogous in many respects, the differences between the two are significant. Unlike out-door positioning systems such as GPS [3], indoor positioning systems cannot depend on high-precision satellite signals [4] or unobstructed sky views. Instead, indoor positioning systems rely on other means of signal propagation, which can include wireless signals (Wi-Fi [5], Bluetooth [6], and RFID [7]), acoustic signals [8], and optical signals (camera-based systems) [9]. However, indoor environments bring to the fore new challenges in the form of increased multipath propagation [10], signal attenuation [11], and environmental changes [12].

The indoor positioning problem has been the subject of extensive research over the past decade, with contributions from a wide range of fields, including communication engineering [13], electrical engineering [14], geography [15],

The associate editor coordinating the review of this manuscript and approving it for publication was Pallab K. Choudhury.

and computer science [16]. The development of accurate indoor positioning systems has the potential to benefit a variety of environments, such as healthcare facilities [17], retail stores [18], universities [19], and airports [20]. For instance, in retail environments, indoor positioning systems could enable targeted advertising [21], personalized promotions [22], and enhance store layout design [23]. In healthcare facilities, such systems could assist in the navigation of patients and healthcare professionals [24], as well as help locate equipment and supplies in real-time [25]. In airports and universities, they could increase travel efficiency [26] and help users navigate complex facilities [27]. Hence, the resolution of the indoor positioning problem is of significant importance and could unlock substantial benefits from location-based services in indoor environments.

Indoor Wi-Fi positioning is becoming an increasingly popular choice for indoor positioning, but currently has many shortcomings. Further research and development are needed to overcome these challenges and optimize indoor Wi-Fi positioning for different indoor environments and use cases. Traditional statistical methods and machine learning models dealing with indoor Wi-Fi localization mainly suffer from the following problems:

1) In indoor Wi-Fi positioning, there may be large differences in Wi-Fi signal strength at different locations, which leads to differences in the transmission time and delay of Wi-Fi signals [28]. In particular, at locations far away from the router, the signal strength will be relatively attenuated, resulting in extended transmission delays of the signal, which in turn creates latency and timing problems.
2) In indoor Wi-Fi positioning, the signal reaching the mobile device may be significantly attenuated and distorted due to factors such as distance, walls and electronic devices [29]. These factors may cause the collected data to contain noise or generate errors.
3) In indoor Wi-Fi localization, the metrics of general methods are usually able to measure only the absolute differences of values in the signals, but cannot reflect information such as their relative positions and trends at different point [30] in time, which leads to the loss of features in the time dimension.

While various techniques have been proposed and utilized in IPS, there is still a need for improved methods that can handle the complexities of indoor environments, such as multipath propagation, signal attenuation, and dynamic environmental changes. To address these challenges, this study focuses on the utilization of the K-nearest neighbor algorithm (KNN) and the distance metric of dynamic time regularization in the domain of indoor Wi-Fi positioning.

KNN offers several advantages in IPS [31], including its simplicity, effectiveness in handling non-linear and non-parametric data, and the ability to adapt to dynamic environmental conditions. Dynamic time regularization, on the other hand, is well-suited for Wi-Fi signals of varying lengths, as it facilitates accurate alignment of sequences and reduces the impact of signal length on distance calculation. Therefore, by combining KNN with dynamic time regularization, we aim to enhance the accuracy and robustness of indoor Wi-Fi positioning systems, especially in the presence of challenges commonly encountered in indoor environments.

To address the challenges mentioned above, this paper proposes a recognition classification strategy for Dynamic Time Warping - K-Nearest Neighbors algorithm (DTW-KNN) embedded with Gaussian sum matrices. The strategy integrates the strengths of Dynamic Time Warping (DTW) [32] and k-Nearest Neighbor (KNN) [33] algorithms to overcome the problems of Wi-Fi signal strength variations across locations, multiple dimension features, and low robustness and accuracy of traditional models. Specifically, KNN is used to calculate the distance between two Wi-Fi signal sequences, where DTW generates a distance matrix to account for the gaps in sequences for KNN algorithm. Moreover, DTW is used to address the polymorphic differences between time sequences with the use of a Gaussian kernel function [34] to solve the distance calculation problem.

The main contributions of this paper are as follows:

First, the combined use of KNN and DTW algorithms can overcome the problem of Wi-Fi signal strength differences across various locations by considering the relationship between Wi-Fi signal transmission delays, signal noise, and distortion. This improves the temporal modeling of Wi-Fi signals and increases the accuracy of the model for localization applications.

Second, the Gaussian kernel function of the DTW algorithm is used instead of the traditional DTW distance calculation method. This function effectively reduces errors and noise caused by signal fading and deformation, thus improving the robustness and accuracy of the localization algorithm. Consequently, more accurate localization of mobile devices in indoor locations is achieved.

Third, the DTW algorithm can capture the polymorphic differences between time series that are not easily measured by conventional methods. By using the DTW algorithm in combination with the KNN algorithm, the accuracy of Wi-Fi signal localization is improved. The temporal characteristics of Wi-Fi signals are better captured, and temporal data in terms of trends and other features are extracted, resulting in a more accurate and robust model.

Overall, the proposed DTW-KNN strategy embedded with Gaussian sum matrices is expected to address the weaknesses of traditional models in indoor Wi-Fi localization. The proposed strategy leverages the strengths of KNN and DTW algorithms and the Gaussian kernel function to significantly improve the accuracy, robustness, and efficiency of the localization process.

The paper is organized as follows: section II presents the theoretical basis of the proposed model in this paper. Subsequently, in Section III, Fast DTW-KNN and fully downsampled DTW-KNN are designed based on the theoretical rationale, and the effects of different distance metric functions are compared. In Section IV, we extend and discuss the experimental results. Section VI gives the conclusions of this study.

II. RELATED WORK

Indoor positioning can be achieved using various technologies, including Wi-Fi [35], Bluetooth Low Energy(BLE) [36], range-free [37], fingerprinting techniques [38], RFID [39], and ultrasound [40]. Each of these technologies has its advantages and limitations [41], and no single technology can provide a universal solution for indoor positioning.

BLE positioning, for example, is suitable for tracking small objects and devices with low power consumption requirements [42], such as indoor navigation for shopping carts and luggage. RFID positioning, on the other hand, can be used for tracking assets within a confined space [43], such as a warehouse [40] or manufacturing plant [44].

Ultrasound positioning relies on the time of flight measurement of acoustic signals [45], which is susceptible to environmental factors and can be affected by ambient noise [46]. Wi-Fi positioning, by contrast, utilizes the signal strength of Wi-Fi access points within the environment and the geometric distance between the user's device and access points to estimate the user's location [47]. This technique can be implemented with existing infrastructure [48] and does not require additional hardware or infrastructure [49], making it a cost-effective solution.

Compared with indoor Wi-Fi positioning technology, both range-free and fingerprinting techniques have their limitations. Range-free technology requires high AP density and no obstructions, otherwise the positioning accuracy will be affected [50]. Fingerprinting technology, on the other hand, requires a complete fingerprint database [51], which is relatively complex. Furthermore, both of these techniques are sensitive to signal fluctuations and external interference which can decrease the positioning accuracy.

In addition, some hybrid technologies are also worthy of our attention. The combination of an IMU (inertial measurement unit) and a camera is a common hybrid technology. The IMU provides acceleration, angular velocity, and orientation information, while the camera provides vision data. This combined technique is useful for applications that need to consider both visual scene and dynamic information. It is widely used in scenarios such as augmented reality (AR), indoor navigation, and virtual reality (VR). However, disadvantages of the IMU+camera [52] approach may include higher requirements for precise calibration and sensitivity to external light conditions. PDR (Gait Recognition) [53] combined with WiFi signal is a commonly used hybrid technique. PDR utilizes on-device inertial sensors such as accelerometers and gyroscopes to estimate the user's gait information and trajectory. By integrating the WiFi signal, especially the signal strength, the positioning accuracy of the PDR system in the indoor environment can be improved. This technology is widely used in areas such as indoor navigation, people tracking and LBS (Location Based Services). However, the disadvantages of PDR include the accumulation of errors when used for a long time, the sensitivity to changes in steps and attitudes, and the inability to provide absolute position information.

Also worth mentioning is UWB (Ultra Wideband) technology [54], which is a wireless communication technology based on short pulses. In indoor positioning, UWB technology has the following characteristics: high precision, high anti-interference, strong refraction transmission ability and trackability. UWB can be used in areas such as indoor people tracking, real-time positioning and object detection. However, the disadvantages of UWB technology include higher equipment and infrastructure costs, and the need for additional deployment and calibration processes.

Compared to other indoor positioning techniques, Wi-Fi positioning has some advantages. For example, it has a larger coverage area [55] and better penetration capabilities [56], which are essential for larger buildings [57] with complex layouts. Wi-Fi signal is also the signal with the simplest requirements on equipment, and it can even be realized with only a mobile phone. Additionally, Wi-Fi signals are less affected by environmental factors [47], making them more reliable for positioning accuracy in real-world scenarios. Wi-Fi positioning can be integrated with a variety of location-based services, such as advertising [58], social networking [59], and search [60], which enables its use in a wide range of applications.

III. MATERIALS AND METHODS

The localization method employed in this study is based on the K-nearest neighbors (KNN) algorithm and dynamic time warping (DTW) [61] technique, with the addition of a Gaussian kernel function for computing the DTW matching distance. Specifically, KNN algorithm is utilized for location determination by measuring the similarity between the available Wi-Fi signals in the training dataset and the test data. DTW is employed as the distance metric in KNN algorithm to effectively handle the high dimensionality and complexity caused by variations in time series such as misalignments and differences in speed. The DTW algorithm generates a distance matrix to handle gaps in the sequences, which is then weighted using a Gaussian kernel function. This approach effectively reduces errors and noise caused by signal attenuation and deformation, thereby improving the robustness and accuracy of the localization algorithm. In summary, our research method integrates the strengths of KNN and DTW algorithms, resulting in improved accuracy, reliability, and real-time capabilities of indoor localization.

A. KNN FOR CLASSIFICATION

K nearest neighbors (KNN) is a supervised classification algorithm that is widely used in classification and regression tasks. The basic principle of KNN is to use the K-nearest neighbors in the training data set that are most similar to the new input sample, and count the frequency of their labels to classify the sample. Specifically, for a given training data set D, where the feature vector x_i and the classification label y_i of each training sample are known, KNN algorithm finds the K nearest training samples by calculating the distance between the input sample u and all the training samples. Then, the algorithm selects a final classification label based on the frequency of the labels of these K training samples.

KNN algorithm can be represented by the following Equation 1:

$$y_u = argmax_{c_j} \sum_i Ind(y_{ij} = c_j) \tag{1}$$

where $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$ represents the training dataset, where x_i is a length-p dimensional vector representing the p-dimensional feature vector of the training sample, and y_i can be a discrete or continuous value label. A classification training set can be represented as $T \subseteq R^p \times Y$, where Y is a set of possible classification labels.

For a new input sample $u \in R^p$, KNN algorithm first calculates the distance between the sample and each sample in the training set, dist(x_i, u), and then finds the K nearest samples $x_{i1}, x_{i2}, \ldots, x_{ik}$ closest to u based on their distance. The algorithm counts the frequency of their labels, and selects the label with the highest frequency, y_u, as the classification result for u.

$$\text{if } y_{ij} = c_j, \ \text{Ind}(y_{ij} = c_j) = 1, \ \text{otherwise} = 0. \tag{2}$$

From Equation 2, it can be seen that the implementation of KNN algorithm mainly involves determining the value of K, calculating the distance, and counting the classification label of K neighbors, and finding the nearest neighbors to the new sample. However, it should be noted that traditional distance measurement methods may not adapt well to the diversity and variation of time series data, and thus KNN algorithm may encounter some problems in some time series-like problems. Therefore, dynamic time warping (DTW) is introduced to improve the performance of KNN algorithm. The algorithm first calculates the DTW distance between the sample to be classified and each sample in the training set to measure their similarity more accurately. Then, select the K training samples with the smallest distance as the nearest neighbors, and make voting decisions according to their categories, and predict the category of the samples to be classified, as shown in Figure 1.

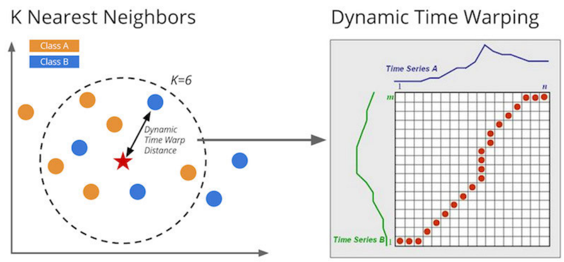

FIGURE 1. DTW distance matrix to calculate the distance between two Wi-Fi sequences.

B. CALCULATE THE DISTANCE MATRIX BETWEEN TWO WI-FI SIGNAL SEQUENCES BY DTW

The DTW algorithm can align signal sequences and calculate the distance matrix between them. Specifically, the DTW algorithm can be divided into the following four steps:

1) CALCULATE THE COST MATRIX

This Equation 3 describes the computation of the cost matrix C by calculating the Euclidean distance between two Wi-Fi signal sequences. The variable N represents the length of the first signal sequence, while M represents the length of the second signal sequence. By calculating the Euclidean distance between each pair of elements, a cost matrix C of size N x M can be obtained. Here, $C(i, j)$ represents the distance between the i-th element of signal sequence A and the j-th element of signal sequence B. The calculation is based on the Euclidean distance formula, which involves subtracting two elements and squaring the result.

First, for two Wi-Fi signal sequences of lengths N and M, respectively, a $N \times M$ cost matrix C can be obtained by calculating the Euclidean distance between them. Here, $C(i, j)$ represents the distance between the i-th element of signal sequence A and the j-th element of signal sequence B, i.e.,

$$C(i, j) = (A(i) - B(j))^2 \tag{3}$$

2) CALCULATE THE ACCUMULATED COST MATRIX

Then, by applying dynamic programming to the cost matrix C, a $N \times M$ accumulated cost matrix D can be obtained. The initial value of the cost matrix C is 0. In Equation 4, $D(i, j)$ represents the minimum cost required to travel from the first element to the i-th element of signal sequence A and from the first element to the j-th element of signal sequence B. The calculation of $D(i, j)$ is as follows:

$$D(i, j) = C(i, j) + \min\{D(i-1, j), D(i, j-1), D(i-1, j-1)\} \tag{4}$$

3) CALCULATE THE DTW DISTANCE MATRIX

Next, a $N \times M$ DTW distance matrix can be obtained by calculating the first row and column of the accumulated cost matrix D, i.e. In the context of the DTW-KNN (Dynamic Time Warping-K Nearest Neighbors) algorithm, the accumulated cost matrix D represents the minimum cost of reaching each element in the cost matrix C.

$$\begin{aligned} &D(1, 1) = C(1, 1) \\ &D(i, 1) = D(i-1, 1) + C(i, 1) \quad \text{for } 2 \leq i \leq N \\ &D(1, j) = D(1, j-1) + C(1, j) \quad \text{for } 2 \leq j \leq M \end{aligned} \tag{5}$$

Here, $D(i, j)$ represents the minimum cost of aligning the first i elements of sequence A with the first j elements of sequence B. It is obtained by accumulating the cost values from previous alignments and considering the cost of the current alignment. The dynamic programming approach used in DTW ensures that the accumulated cost matrix D represents the sum of the minimum costs for all possible alignments.

4) CALCULATE THE DTW DISTANCE

Finally, the DTW distance between two Wi-Fi signal sequences can be calculated by flattening the DTW distance matrix D into a vector, as shown in Figure 2. Specifically, all the elements of D can be flattened into a length-$N \times M$ vector d_{DTW}, which can then be used to calculate the Euclidean distance between d_{DTW} (Equation 6) and the same-length vector $d_{A,B}$, where $d_{A,B}(i,j)$ is a $N \times M$ matrix representing the Euclidean distance between the i-th element of signal sequence A and the j-th element of signal sequence B, i.e.,

$$d_{A,B}(i,j) = \sqrt{(A(i) - B(j))^2} \tag{6}$$

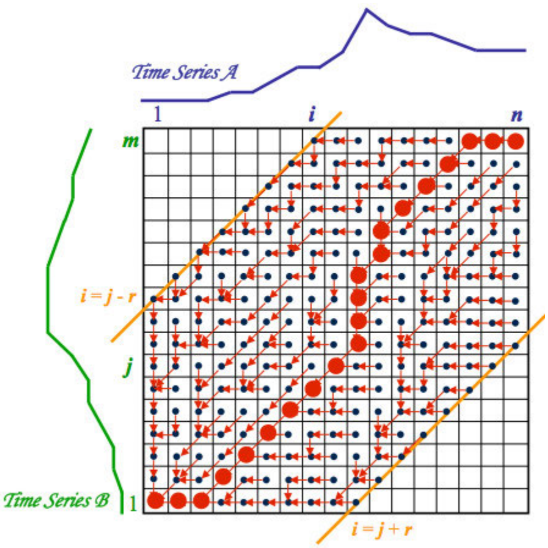

FIGURE 2. The backtracking process of DTW distance matrix calculation distance.

C. USE GAUSSIAN KERNEL FUNCTION TO CALCULATE THE DISTANCE OF DTW

Using Manhattan distances to calculate the DTW distance matrix is a common practice, but has many drawbacks. One of them is that the Manhattan distance only considers the simple distance between points on the x and y axes, and does not consider the connection between the points. In contrast, indoor Wi-Fi signal localization precisely uses the significant relationship between individual sequences for distance calculation. Another point is that in the case of large-scale DTW deformation, the Manhattan distance may not correctly match the true overlapping part of the deformed signal, which leads to a decrease in localization accuracy. In contrast, using Gaussian kernel function to calculate the DTW distance matrix can better balance local similarity and deformation features, and thus is more suitable for complex indoor localization scenarios. Specifically, the following Equation 7 based on the Gaussian kernel function can be used to calculate the DTW distance:

$$d_{DTW}(i,j) = exp\left(\frac{(A(i,j) - A(i-1,j))^2 + (A(i,j) - A(i,j-1))^2}{\sigma^2}\right) \tag{7}$$

Here, $d_{DTW}(i,j)$ is the value at position (i,j) in the DTW distance matrix, $A(i,j)$ is the cost matrix in the DTW algorithm, and σ is the parameter of the Gaussian kernel function, which determines the width of the Gaussian distribution.

By using this method, the DTW distance matrix can be transformed into a kernel matrix and then the kernel method can be used for classification The kernel method is a technique for linearizing a nonlinear problem by transforming the original data into a high-dimensional feature space, making the Wi-Fi signal data linearly separable in the new feature space.

IV. EXPERIMENTS

To highlight the effectiveness of the model, We experimented with the commonly used Manhattan distance and our proposed Gaussian kernel function in the distance metric of the DTW algorithm, respectively. And the parameters are optimized and adjusted for them.

Meanwhile, in order to cope with the sparsity and continuity problems of Wi-Fi data in different scenarios, We have also made a full classification discussion in feature selection. On the one hand, when targeting the data with good continuity, feature selection and classification can be performed quickly by intermittent downsampling, as shown in Section IV-C. On the other hand, when targeting the more sparse cases, the full data was directly used for testing, as shown in Section IV-D.

A. DATA SET

To test the effectiveness of our model, We use the Wireless Indoor Localization Data Set [62], a publicly available indoor localization dataset used to study the performance of different indoor localization algorithms. The dataset is published by the UCI Machine Learning Repository's machine learning website.Wireless Indoor Localization Data Set contains information about the Wi-Fi signal strength in different indoor rooms.

The dataset contains seven values representing the Wi-Fi signal strength emitted by seven routers at the current location. Each sample in the dataset also includes a label, which indicates the indoor room in which the sample is located.

Using this dataset to assist researchers in their experiments can be a good way to test the performance of different types of algorithms, and the dataset is divided into a training set and a test set according to 8:2.

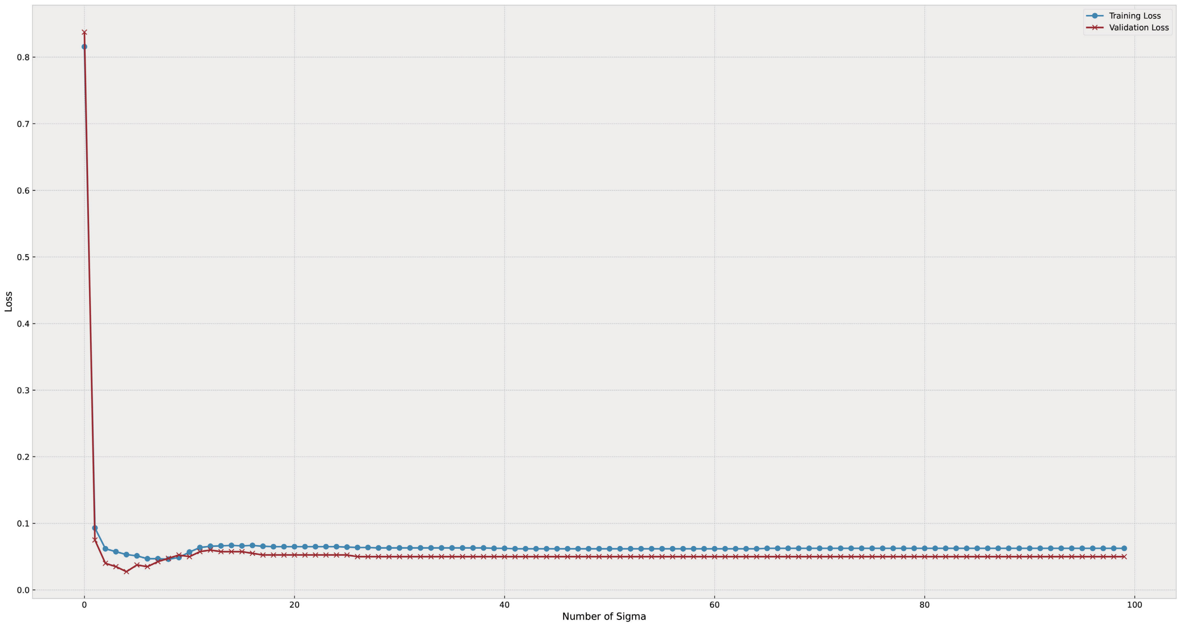

FIGURE 3. Correspondence between hyperparameter sigma and loss when the interval is 1 unit.

B. HYPERPARAMETER OPTIMIZATION AND SELECTION OF DTW-KNN MODEL BASED ON GAUSSIAN KERNEL MATRIX

1) SELECTION OF SIGMA HYPERPARAMETERS IN GAUSSIAN KERNEL FUNCTION

Optimization and selection of sigma hyperparameters in Gaussian kernel function In practical applications, we usually hope to improve the performance of the model by optimizing the hyperparameters of the model. In the distance measure of DTW, we can use the Gaussian kernel function to weight the sample distance, so that the samples with far distances have less influence on the prediction results. A common parameter of the Gaussian kernel function is the sigma value, which represents the standard deviation of the Gaussian distribution curve and is used to control the weight distribution ratio of the distance.

In order to optimize and select the most suitable hyperparameters for indoor Wi-Fi scenarios, the sigma value setting mainly includes two conditions: the first, the average value of the data sample interval; the second, a multiple of the standard deviation of the data distribution. Therefore, we first carried out the primary selection of sigma to determine the approximate range of the optimal solution. That is, by traversing all values of sigma between 0 and 100, and then recording the loss corresponding to the model, as shown in Figure 3. It can be seen that the loss corresponding to sigma with a value between 0 and 10 is the smallest. Therefore, the optimal sigma value should be between 0 and 10. Based on this result, I conduct further experiments.

It can be seen from the Figure 4 that because the sigma value is too large or too small, the loss of the verification set is also relatively large. This is because when the sigma value is too large, the specified distance distribution affects the data points too uniformly, making it impossible to effectively distinguish the data points with a closer distance from the data points with a longer distance, thereby affecting the classification effect of the model. Conversely, if the sigma value is too small, too much noise may be introduced, making the model prone to overfitting. It can be seen that when the sigma value is 4, the loss of the validation set is the smallest and the model performs best.

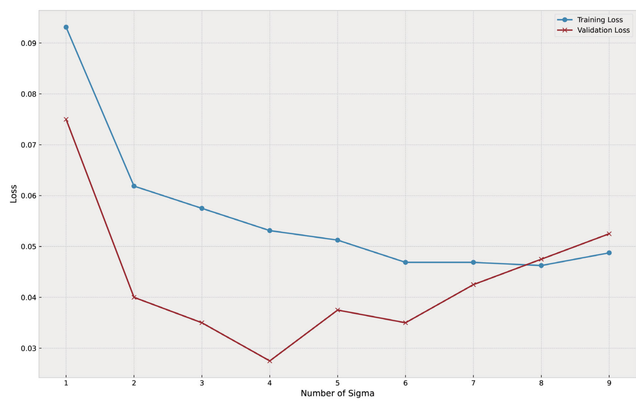

FIGURE 4. The change of the verification set loss curve with the value of sigma from 1 to 9 after applying the Gaussian kernel function.

2) SELECTION OF NEIGHBOR DISTANCE IN K-NEAREST NEIGHBOR MODEL

When determining the range of neighbor distance, my idea is to conduct data exploration and visual analysis first, deter-

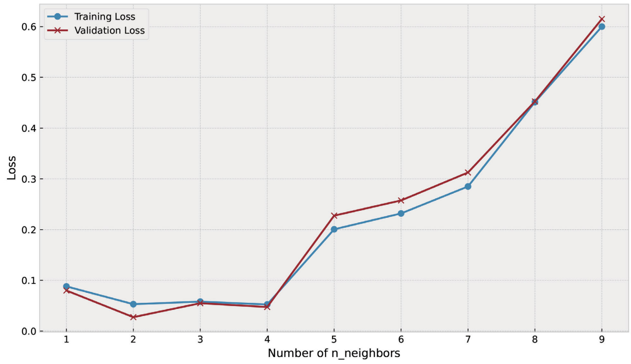

FIGURE 5. Variation of the validation set loss curve with the value of n_neighbors from 1 to 9.

mine a rough distance range according to the distribution of data and task requirements, and then gradually refine and adjust it according to the experimental results. In practical applications, the common neighbor distance ranges from 1 to 10. Figure 5 presents the performance of the model with n_neighbors between 0 and 10. It can be seen that when n_neighbors is equal to 2, the loss of the verification set reaches the minimum and the model performs best.

C. INTERMITTENT DOWNSAMPLING OF DTW-KNN FOR APPLICATIONS

High noise and continuity issues often occur when using Wi-Fi signal strength for positioning. In order to solve this problem, by default, each sample in the dataset will be down-sampled on the basis of every 100 data, that is, every 100 points apart when sampling, a data input model is taken to compress the data and make the dataset more compact. Readability and interpretability.

1) FAST DTW-KNN RESULTS USING MANHATTAN DISTANCE

Faster DTW-KNN results using Manhattan distance, faster than without downsampling, are shown in the table 1, which we call Fast DTW-KNN (FDK). The FDK confusion matrix is shown in Figure 6.

Specifically, the data in each sample are taken from the RSSI readings of Wi-Fi devices and are sampled every 100 readings at the time of sampling in order to reduce the amount of data in the sample, while also retaining the representativeness of the sample data for subsequent processing and analysis.

2) FAST DTW-KNN RESULTS USING GAUSSIAN KERNEL MATRIX

In this experiment, based on the indoor Wi-Fi positioning data set, the Fast DTW-KNN algorithm using the Gaussian kernel matrix was tested, and the model performance was also evaluated. The results are shown in Table 2. Figure 7 is the corresponding confusion.

TABLE 1. FDK results for Manhattan distance.

	precision	recall	f1-score	support
Room1	96.3%	97.2%	96.7%	106
Room2	90.3%	98.8%	94.4%	85
Room3	98.0%	86.8%	92.1%	114
Room4	94.9%	98.9%	96.9%	95
accuracy		95.0%		400
macro avg	94.9%	95.4%	95.0%	400
weighted avg	95.2%	95.0%	94.9%	400

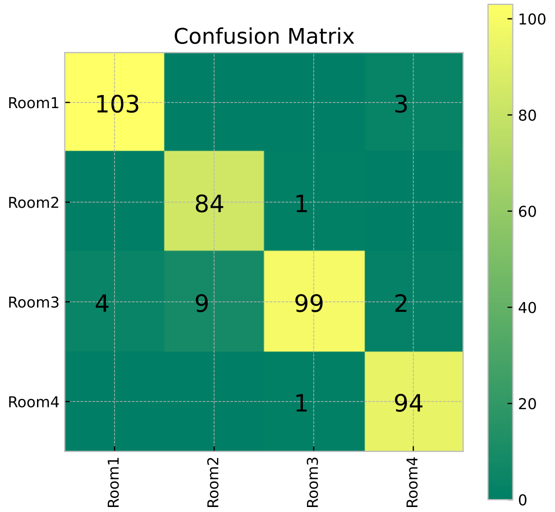

FIGURE 6. FDK confusion matrix using Manhattan distance.

TABLE 2. FDK results for Gaussian kernel matrix.

	precision	recall	f1-score	support
Room1	97.2%	99.0%	98.1%	105
Room2	91.4%	97.7%	94.4%	87
Room3	97.0%	86.7%	91.6%	113
Room4	94.9%	98.9%	96.9%	95
accuracy		95.3%		400
macro avg	95.1%	95.6%	95.3%	400
weighted avg	95.4%	95.3%	95.2%	400

In order to see the classification of each room more intuitively, I drew a scatter plot of the real value and the predicted value, as shown in Figure 8. The value on the vertical axis represents the number of the room category, and the horizontal axis represents the value of the sample. It can be seen that only room 3 has a certain amount of misclassification, and rooms 1, 2, and 4 have almost none. Intuitively, the classification effect of FDK using Gaussian and matrix can be seen that most rooms can be accurately distinguished.

3) DISCUSSION OF FAST DTW-KNN FOR DIFFERENT DISTANCE METRICS

From the results of the two experiments, the following conclusions can be drawn:

Gaussian kernel function distance calculation is slightly better than Manhattan distance calculation. By comparing the results of the two experiments, it can be found that the

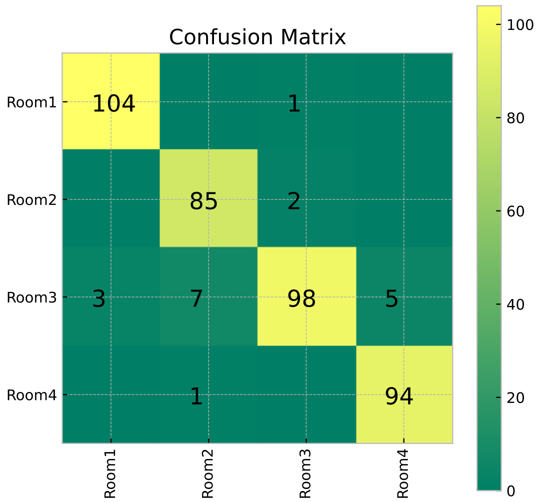

FIGURE 7. FDK confusion matrix using Gaussian kernel matrix.

Gaussian kernel function distance calculation is better than the Manhattan distance calculation in terms of precision, recall and f1-score. For example, in Room1, the precision rate and recall rate calculated by Gaussian kernel function distance are higher than those calculated by Manhattan distance.

There is a certain degree of volatility in the evaluation indicators. By comparing the weighted average of the two experimental results, it can be seen that there is a certain degree of volatility in the performance of the evaluation indicators. For example, in the first experiment, the weighted avg indicator is slightly lower than 95%; while in the second experiment, the weighted avg indicator is slightly higher than 95%. This shows that the performance of indoor Wi-Fi positioning algorithms may be affected by different data sets and experimental conditions, and needs to be optimized and improved in practical applications.

In summary, the experimental results obtained by Gaussian kernel function distance calculation are relatively better than those calculated by Manhattan distance. FDK has good performance and generalization ability in indoor Wi-Fi positioning, and can be adapted to the application scenario of fast analysis, but the prediction effect in different rooms is different.

D. DTW-KNN USING ALL SAMPLES

In Wi-Fi signal strength localization, the process of data acquisition often has high noise and data sparsity problems. In order to retain more data information, no downsampling process was performed in the dataset, and the entire data was directly used for acquisition and storage. Each sample in the dataset contains the Wi-Fi signal strength values emitted by seven routers in the indoor environment, and the information of the room where each sample is located is labeled.

1) DTW-KNN RESULTS USING MANHATTAN DISTANCE

More useful information and more comprehensive data characteristics can be obtained by using all the data, and the results are shown in Table 3. Figure 9 is the corresponding confusion.

TABLE 3. DK results for Manhattan distance.

	precision	recall	f1-score	support
Room1	100.0%	100.0%	100.0%	107
Room2	94.6%	91.7%	93.1%	96
Room3	92.1%	94.9%	93.5%	98
Room4	100.0%	100.0%	100.0%	99
accuracy		96.8%		400
macro avg	96.7%	96.6%	96.6%	400
weighted avg	96.8%	96.8%	96.7%	400

TABLE 4. DK results for Gaussian kernel matrix.

	precision	recall	f1-score	support
Room1	100.0%	100.0%	100.0%	107
Room2	97.8%	94.8%	96.3%	96
Room3	94.1%	97.9%	96.0%	97
Room4	100.0%	99.0%	99.5%	100
accuracy		98.0%		400
macro avg	98.0%	97.9%	97.9%	400
weighted avg	98.0%	98.0%	98.0%	400

2) DTW-KNN RESULTS USING GAUSSIAN KERNEL MATRIX

In this experiment, based on the indoor Wi-Fi positioning data set, the DTW-KNN algorithm using the Gaussian kernel matrix was tested, and the model performance was also evaluated. The results are shown in Table 4. Figure 10 is the corresponding confusion.

In order to see the classification of each room more intuitively, I drew a scatter plot of the real value and the predicted value, as shown in Figure 11. The value on the vertical axis represents the number of room categories, and the horizontal axis represents the value of the sample. It can be seen that no room has too many misclassified samples. Among them, the classification of Room 4 is all correct, and there is no misclassified sample. In rooms 1, 2, and 3, there are only sporadic misclassified samples that are not concentrated. The DK model is the one with the best indoor Wi-Fi signal classification among all the models tested in this paper.

3) DISCUSSION OF DTW-KNN FOR DIFFERENT DISTANCE METRICS

From Tables 1, 2, 3, and 4, it can be seen that the accuracy of the indoor Wi-Fi localization algorithm derived from both the Manhattan distance calculation and the Gaussian kernel function calculation without downsampling is higher than the accuracy of the previously downsampled data.

The experimental results without downsampling data are more accurate. Since the downsampling operation reduces the amount of data, which may lead to some degree of information loss, the model is able to use more information to improve the accuracy in the experiments without downsampling.

The indicator fluctuations in these two experiments are smaller than those in the FDK experiments, and this change can be seen through the weighted avg indicator.

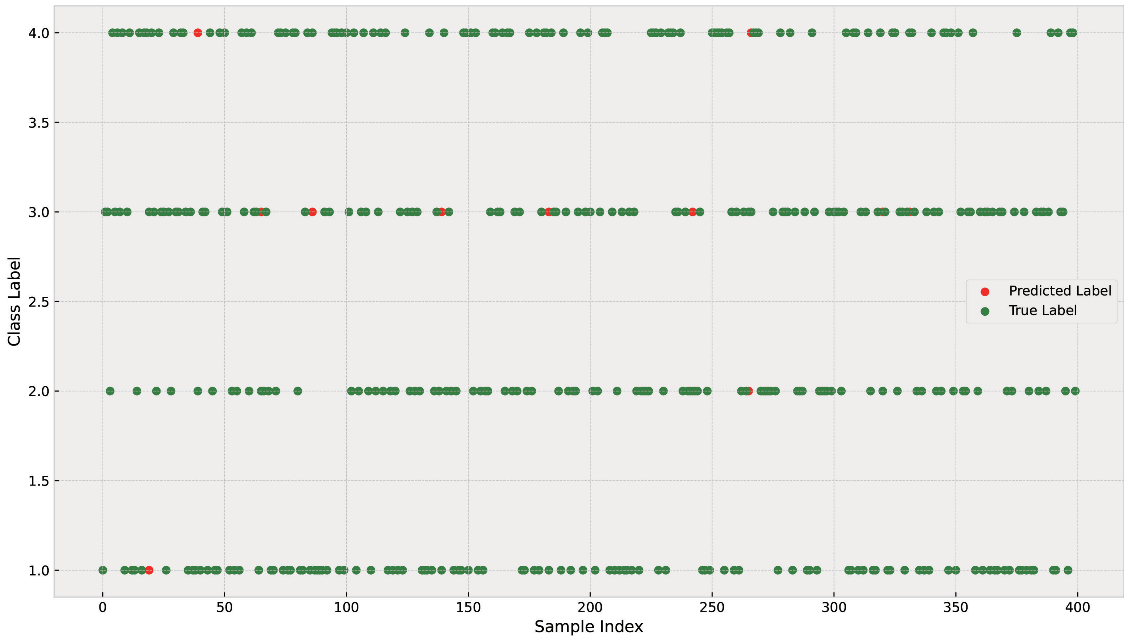

FIGURE 8. Scatter plot of true and predicted values for FDK model using Gaussian and matrix.

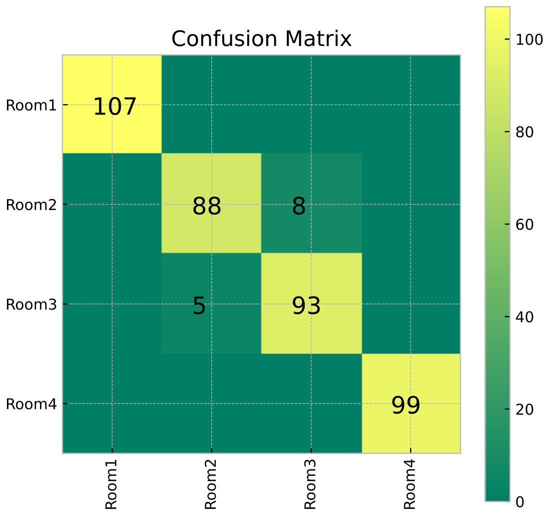

FIGURE 9. DK confusion matrix using Manhattan distance.

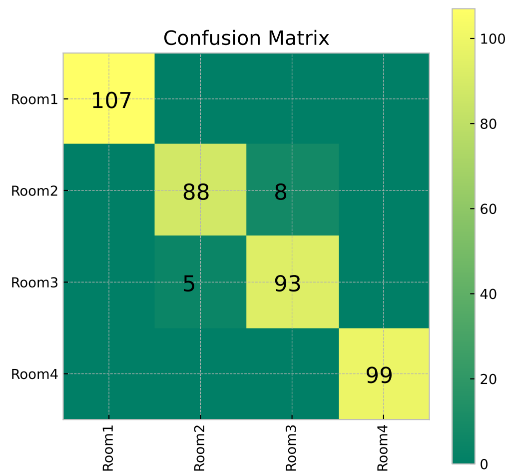

FIGURE 10. FDK confusion matrix using Gaussian kernel matrix.

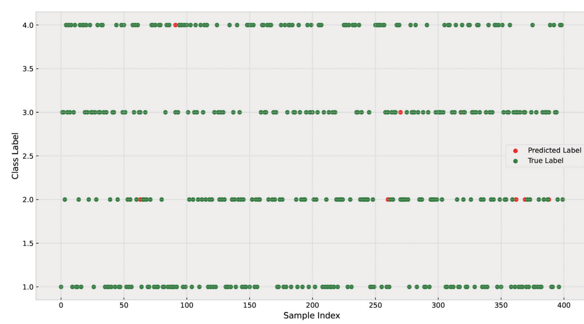

FIGURE 11. Scatter plot of true and predicted values for DK model using Gaussian and matrix.

V. RESULT AND DISCUSSION

In order to show the performance comparison of different algorithms [63] more intuitively and make our algorithm more convincing, we made the experimental results into a visual bar graph, as shown in Figure 12.

As can be seen from Figure 12, our proposed algorithm achieves good accuracy performance under different distance calculation methods, and only FDK using Manhattan distance is slightly lower than FPSOGSA-NN compared with other algorithms. both our proposed FDK using Gaussian kernel function and DK have greater advantages.

In addition, the accuracy performance of FDK algorithm and DK algorithm is relatively smooth under different distance calculation methods. In the field of indoor Wi-Fi positioning, FDK algorithm and DK algorithm perform well

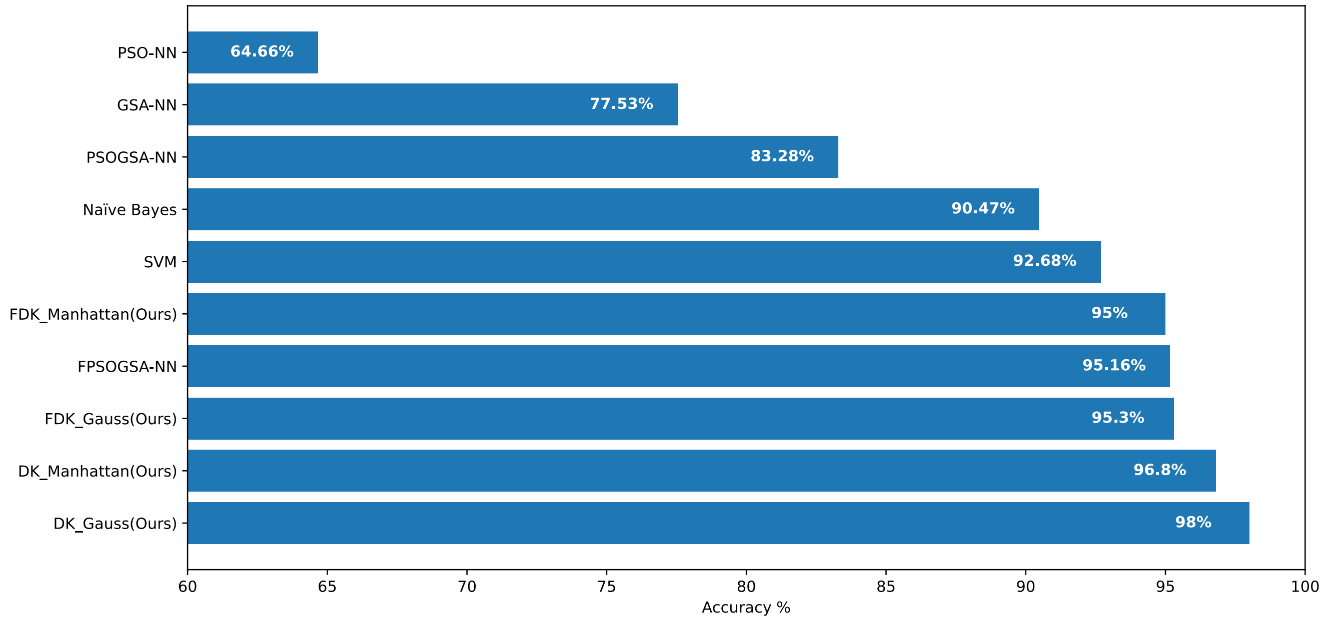

FIGURE 12. Comparison of the accuracy of FDK and DK with other indoor Wi-Fi positioning algorithms [63].

after parameters such as distance calculation method and constraint window length are optimized, and can be used as An effective localization algorithm. At the same time, they also have advantages in some specific scenarios, which can expand the selection range of algorithms.

Although this study proposes some ideas for algorithm selection and optimization from the experimental results and analysis, there are still some limitations and deficiencies. First, we only conducted experimental tests on the dataset and did not verify it on data in other real scenarios, so the applicability and performance of the algorithm may vary. In addition, in the setting of algorithm parameters, we did not optimize and debug the system, which may also affect the performance of the algorithm.

Future research directions can be carried out from the following aspects:

1) Expand datasets and test scenarios: test and verify the applicability and performance of algorithms for different indoor scenarios and datasets.;
2) Meticulous parameter tuning: For algorithms that require parameter optimization, systematically optimize and debug to improve the performance and stability of the algorithm.;
3) Combining machine learning methods: This study only uses relevant strategies of machine learning. If methods such as deep learning are used, more accurate and robust positioning models and algorithms may be established.

VI. CONCLUSION

In this paper, we design a DTW-KNN model based on Gaussian kernel function, and discuss the application and optimization method of this model in the field of indoor Wi-Fi positioning. Through the analysis of experimental data, we found that the distance calculation method has an important impact on the performance, accuracy and robustness of the algorithm. In the comparison of distance calculation methods, the indoor Wi-Fi positioning algorithm calculated by the Gaussian kernel function performs better than the Manhattan distance, and has better robustness and more flexible calculation properties.

At the same time, we also compared the performance gap between fast DTW-KNN (FKD) using downsampling and DTW-KNN (DK) with full data, and tested it on public datasets. In real-time and continuous data, it is recommended to consider FDK; in the case of high accuracy requirements and discrete data, it is recommended to consider DK; in practical applications, we need to consider specific scenarios and data sets Select the appropriate distance calculation method and algorithm to achieve better accuracy and performance. In summary, using Gaussian kernel function to optimize the distance metric of DTW-KNN is scientifically reasonable and helps to improve the accuracy and efficiency of indoor positioning technology.

REFERENCES

[1] Z. Jiao, L. Chen, X. Lu, Z. Liu, X. Zhou, Y. Zhuang, and G. Guo, "Carrier phase ranging with DTMB signals for urban pedestrian localization and GNSS aiding," *Remote Sens.*, vol. 15, no. 2, p. 423, Jan. 2023.

[2] C. Y. T. Kwok, M. S. Wong, S. Griffiths, F. Y. Y. Wong, R. Kam, D. C. W. Chin, G. Xiong, and E. Mok, "Performance evaluation of iBeacon deployment for location-based services in physical learning spaces," *Appl. Sci.*, vol. 10, no. 20, p. 7126, Oct. 2020.

[3] S. Nilwong, D. Hossain, S.-I. Kaneko, and G. Capi, "Deep learning-based landmark detection for mobile robot outdoor localization," *Machines*, vol. 7, no. 2, p. 25, Apr. 2019.

[4] J. Gante, L. Sousa, and G. Falcao, "Dethroning GPS: Low-power accurate 5G positioning systems using machine learning," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 10, no. 2, pp. 240–252, Jun. 2020.

[5] F. Qin, T. Zuo, and X. Wang, "CCpos: WiFi fingerprint indoor positioning system based on CDAE-CNN," *Sensors*, vol. 21, no. 4, p. 1114, Feb. 2021.

[6] X. Li, J. Wang, and C. Liu, "A Bluetooth/PDR integration algorithm for an indoor positioning system," *Sensors*, vol. 15, no. 10, pp. 24862–24885, Sep. 2015.

[7] Y.-F. Hsu, C.-S. Cheng, and W.-C. Chu, "COMPASS: An active RFID-based real-time indoor positioning system," *Human-Centric Comput. Inf. Sci.*, vol. 12, no. 7, pp. 1–20, 2022.

[8] J. Kunhoth, A. Karkar, S. Al-Maadeed, and A. Al-Ali, "Indoor positioning and wayfinding systems: A survey," *Human-Centric Comput. Inf. Sci.*, vol. 10, no. 1, pp. 1–41, Dec. 2020.

[9] S. Bian, P. Hevesi, L. Christensen, and P. Lukowicz, "Induced magnetic field-based indoor positioning system for underwater environments," *Sensors*, vol. 21, no. 6, p. 2218, Mar. 2021.

[10] E. Sippel, J. Geiss, S. Brückner, P. Gröschel, M. Hehn, and M. Vossiek, "Exchanging bandwidth with aperture size in wireless indoor localization–or why 5G/6G systems with antenna arrays can outperform UWB solutions," *IEEE Open J. Veh. Technol.*, vol. 2, pp. 207–217, 2021.

[11] T. Kim Geok, K. Zar Aung, M. Sandar Aung, M. Thu Soe, A. Abdaziz, C. Pao Liew, F. Hossain, C. P. Tso, and W. H. Yong, "Review of indoor positioning: Radio wave technology," *Appl. Sci.*, vol. 11, no. 1, p. 279, Dec. 2020.

[12] E. R. Magsino, I. W. Ho, and Z. Situ, "The effects of dynamic environment on channel frequency response-based indoor positioning," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–6.

[13] V. Renaudin, "Evaluating indoor positioning systems in a shopping mall: The lessons learned from the IPIN 2018 competition," *IEEE Access*, vol. 7, pp. 148594–148628, 2019.

[14] B. AL-Madani, F. Orujov, R. Maskeliūnas, R. Damaševičius, and A. Venčkauskas, "Fuzzy logic type-2 based wireless indoor localization system for navigation of visually impaired people in buildings," *Sensors*, vol. 19, no. 9, p. 2114, May 2019.

[15] Y. Liu, Q. Zhu, F. Cao, J. Chen, and G. Lu, "High-resolution remote sensing image segmentation framework based on attention mechanism and adaptive weighting," *ISPRS Int. J. Geo-Inf.*, vol. 10, no. 4, p. 241, Apr. 2021. [Online]. Available: https://www.mdpi.com/2220-9964/10/4/241

[16] J. Chen, Y. Liu, and Q. Zhu, "Enterprise profitability and financial evaluation model based on statistical modeling: Taking tencent music as an example," *Mathematics*, vol. 10, no. 12, p. 2107, Jun. 2022.

[17] T. Å. Evjen, S. R. H. Raviz, S. A. Petersen, and J. Krogstie, "Smart facility management: Future healthcare organization through indoor positioning systems in the light of enterprise BIM," *Smart Cities*, vol. 3, no. 3, pp. 793–805, Aug. 2020.

[18] H. Hwangbo, J. Kim, Z. Lee, and S. Kim, "Store layout optimization using indoor positioning system," *Int. J. Distrib. Sensor Netw.*, vol. 13, no. 2, Feb. 2017, Art. no. 1550147717692585.

[19] I. Cil, F. Arisoy, E. Özgürbüz, A. Y. Cil, and H. Kilinç, "Indoor positioning technology selection using a combined AHP and PROMETHEE method at SEDEF shipyard," *J. ETA Maritime Sci.*, vol. 10, no. 2, pp. 108–123, Jun. 2022.

[20] P. Singh, M. Jain, L. Rawal, and P. Nagrath, "iBeacon-based indoor positioning systems for airports," *Int. J. Comput. Appl.*, vol. 975, p. 8887, Jan. 2015.

[21] W. Shule, C. M. Almansa, J. P. Queralta, Z. Zou, and T. Westerlund, "UWB-based localization for multi-UAV systems and collaborative heterogeneous multi-robot systems," *Proc. Comput. Sci.*, vol. 175, pp. 357–364, 2020.

[22] N. Đurević, A. Labus, D. Barać, M. Radenković, and M. Despotović-Zrakić, "An approach to assessing shopper acceptance of beacon triggered promotions in smart retail," *Sustainability*, vol. 14, no. 6, p. 3256, Mar. 2022.

[23] K. Kan, Z. Yang, P. Lyu, Y. Zheng, and L. Shen, "Numerical study of turbulent flow past a rotating axial-flow pump based on a level-set immersed boundary method," *Renew. Energy*, vol. 168, pp. 960–971, May 2021.

[24] M. Patnaik and S. Mishra, "Indoor positioning system assisted big data analytics in smart healthcare," in *Connected e-Health: Integrated IoT and Cloud Computing*. Cham, Switzerland: Springer, 2022, pp. 393–415.

[25] N. Soltanmohammadlou, S. Sadeghi, C. K. H. Hon, and F. Mokhtarpour-Khanghah, "Real-time locating systems and safety in construction sites: A literature review," *Saf. Sci.*, vol. 117, pp. 229–242, Aug. 2019.

[26] Q. Zhu, Y. Liu, M. Liu, S. Zhang, G. Chen, and H. Meng, "Intelligent planning and research on urban traffic congestion," *Future Internet*, vol. 13, no. 11, p. 284, Nov. 2021.

[27] K. Kan, H. Chen, Y. Zheng, D. Zhou, M. Binama, and J. Dai, "Transient characteristics during power-off process in a shaft extension tubular pump by using a suitable numerical model," *Renew. Energy*, vol. 164, pp. 109–121, Feb. 2021.

[28] S. Lee, J. Kim, and N. Moon, "Random forest and WiFi fingerprint-based indoor location recognition system using smart watch," *Human-Centric Comput. Inf. Sci.*, vol. 9, no. 1, pp. 1–14, Dec. 2019.

[29] K. A. Nguyen, Z. Luo, G. Li, and C. Watkins, "A review of smartphones-based indoor positioning: Challenges and applications," *IET Cyber-Syst. Robot.*, vol. 3, no. 1, pp. 1–30, Mar. 2021.

[30] Y. Ge, A. Taha, S. A. Shah, K. Dashtipour, S. Zhu, J. Cooper, Q. H. Abbasi, and M. A. Imran, "Contactless WiFi sensing and monitoring for future healthcare–emerging trends, challenges, and opportunities," *IEEE Rev. Biomed. Eng.*, vol. 16, pp. 171–191, 2023.

[31] A. Poulose and D. S. Han, "Hybrid deep learning model based indoor positioning using Wi-Fi RSSI heat maps for autonomous applications," *Electronics*, vol. 10, no. 1, p. 2, Dec. 2020. [Online]. Available: https://www.mdpi.com/2079-9292/10/1/2

[32] J. Chen, S. Song, Y. Gong, and S. Zhang, "An indoor fusion navigation algorithm using HV-derivative dynamic time warping and the chicken particle filter," *Satell. Navigat.*, vol. 3, no. 1, p. 13, Dec. 2022.

[33] A. Sobehy, É. Renault, and P. Mühlethaler, "CSI-MIMO: K-nearest neighbor applied to indoor localization," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[34] M. Zhou, X. Li, Y. Wang, S. Li, Y. Ding, and W. Nie, "6G multisource-information-fusion-based indoor positioning via Gaussian kernel density estimation," *IEEE Internet Things J.*, vol. 8, no. 20, pp. 15117–15125, Oct. 2021.

[35] L. Huang, B. Yu, H. Li, H. Zhang, S. Li, R. Zhu, and Y. Li, "HPIPS: A high-precision indoor pedestrian positioning system fusing WiFi-RTT, MEMS, and map information," *Sensors*, vol. 20, no. 23, p. 6795, Nov. 2020.

[36] P. Bencak, D. Hercog, and T. Lerher, "Indoor positioning system based on Bluetooth low energy technology and a nature-inspired optimization algorithm," *Electronics*, vol. 11, no. 3, p. 308, Jan. 2022.

[37] E. Shakshuki, A. A. Elkhail, I. Nemer, M. Adam, and T. Sheltami, "Comparative study on range free localization algorithms," *Proc. Comput. Sci.*, vol. 151, pp. 501–510, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050919305307

[38] A. Poulose and D. S. Han, "Performance analysis of fingerprint matching algorithms for indoor localization," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, Feb. 2020, pp. 661–665.

[39] I. Tabakh, S. Das, M. Jorio, N. E. A. El Idrissi, S. Mohapatra, and D. Barad, "Defected ground structure (DGS) incorporated RFID reader antenna array for indoor positioning systems at 2.45 GHz," *Int. J. Microw. Opt. Technol.*, vol. 15, no. 6, pp. 517–524, 2020.

[40] K. Kan, Z. Xu, H. Chen, H. Xu, Y. Zheng, D. Zhou, A. Muhirwa, and B. Maxime, "Energy loss mechanisms of transition from pump mode to turbine mode of an axial-flow pump under bidirectional conditions," *Energy*, vol. 257, Oct. 2022, Art. no. 124630.

[41] H. Obeidat, W. Shuaieb, O. Obeidat, and R. Abd-Alhameed, "A review of indoor localization techniques and wireless technologies," *Wireless Pers. Commun.*, vol. 119, no. 1, pp. 289–327, Jul. 2021.

[42] S. Subedi and J.-Y. Pyun, "A survey of smartphone-based indoor positioning system using RF-based wireless technologies," *Sensors*, vol. 20, no. 24, p. 7230, Dec. 2020.

[43] J. Park, H. Kim, J. Yoon, H. Kim, C. Park, and D. Hong, "Development of an ultrasound technology-based indoor-location monitoring service system for worker safety in shipbuilding and offshore industry," *Processes*, vol. 9, no. 2, p. 304, Feb. 2021.

[44] F. Liu, J. Liu, Y. Yin, W. Wang, D. Hu, P. Chen, and Q. Niu, "Survey on WiFi-based indoor positioning techniques," *IET Commun.*, vol. 14, no. 9, pp. 1372–1383, Jun. 2020.

[45] S. A. Pullano, M. G. Bianco, D. C. Critello, M. Menniti, A. La Gatta, and A. S. Fiorillo, "A recursive algorithm for indoor positioning using pulse-echo ultrasonic signals," *Sensors*, vol. 20, no. 18, p. 5042, Sep. 2020.

[46] D. Delabie, B. Cox, L. De Strycker, and L. Van der Perre, ''Techtile: A flexible testbed for distributed acoustic indoor positioning and sensing,'' in *Proc. IEEE Sensors Appl. Symp. (SAS)*, Aug. 2022, pp. 1–6.
[47] D. B. Ninh, J. He, V. T. Trung, and D. P. Huy, ''An effective random statistical method for indoor positioning system using WiFi fingerprinting,'' *Future Gener. Comput. Syst.*, vol. 109, pp. 238–248, Aug. 2020.
[48] A. Haider, Y. Wei, S. Liu, and S.-H. Hwang, ''Pre- and post-processing algorithms with deep learning classifier for Wi-Fi fingerprint-based indoor positioning,'' *Electronics*, vol. 8, no. 2, p. 195, Feb. 2019.
[49] Y. Liu, G. Lu, J. Chen, and Q. Zhu, ''Exploration of internal and external factors of swimmers' performance based on biofluid mechanics and computer simulation,'' *Int. J. Environ. Res. Public Health*, vol. 18, no. 12, p. 6471, Jun. 2021.
[50] S. Ghorpade, M. Zennaro, and B. Chaudhari, ''Survey of localization for Internet of Things nodes: Approaches, challenges and open issues,'' *Future Internet*, vol. 13, no. 8, p. 210, Aug. 2021.
[51] J. M. Rocamora, I. W.-H. Ho, W. Mak, and A. P. Lau, ''Survey of CSI fingerprinting-based indoor positioning and mobility tracking systems,'' *IET Signal Process.*, vol. 14, no. 7, pp. 407–419, Sep. 2020.
[52] S. Wen, Z. Ge, D. Yuan, Y. Chen, F. Wen, J. Xu, and W. Guan, ''Enhanced pedestrian navigation on smartphones with VLP-assisted PDR integration,'' *IEEE Sensors J.*, vol. 23, no. 14, pp. 15952–15962, Jul. 2023.
[53] A. Poulose, J. Kim, and D. S. Han, ''A sensor fusion framework for indoor localization using smartphone sensors and Wi-Fi RSSI measurements,'' *Appl. Sci.*, vol. 9, no. 20, p. 4379, Oct. 2019.
[54] A. Poulose and D. S. Han, ''Feature-based deep LSTM network for indoor localization using UWB measurements,'' in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, Apr. 2021, pp. 298–301.
[55] Z. Li, X. Zhao, F. Hu, Z. Zhao, J. L. C. Villacrés, and T. Braun, ''SoiCP: A seamless outdoor–indoor crowdsensing positioning system,'' *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8626–8644, Oct. 2019.
[56] K. Kim, S. Li, M. Heydariaan, N. Smaoui, O. Gnawali, W. Suh, M. J. Suh, and J. I. Kim, ''Feasibility of LoRA for smart home indoor localization,'' *Appl. Sci.*, vol. 11, no. 1, p. 415, Jan. 2021.
[57] K. Kan, M. Binama, H. Chen, Y. Zheng, D. Zhou, W. Su, and A. Muhirwa, ''Pump as turbine cavitation performance for both conventional and reverse operating modes: A review,'' *Renew. Sustain. Energy Rev.*, vol. 168, Oct. 2022, Art. no. 112786.
[58] Y. Bae and D. Shin, ''BLE-based indoor positioning using extended advertisement,'' in *Proc. 37th Int. Tech. Conf. Circuits/Syst., Comput. Commun. (ITC-CSCC)*, Jul. 2022, pp. 667–670.
[59] T. Ahmad, X. J. Li, B.-C. Seet, and J.-C. Cano, ''Social network analysis based localization technique with clustered closeness centrality for 3D wireless sensor networks,'' *Electronics*, vol. 9, no. 5, p. 738, Apr. 2020.
[60] Y. Cui, Y. Zhang, Z. Wang, and H. Fu, ''Integrated WiFi/MEMS indoor navigation based on searching space limiting and self-calibration,'' *Arabian J. Sci. Eng.*, vol. 45, no. 4, pp. 3015–3024, Apr. 2020.
[61] Z. Shi, Z. Zhang, Y. Shu, P. Cheng, and J. Chen, ''Indoor navigation leveraging gradient WiFi signals,'' in *Proc. 15th ACM Conf. Embedded Netw. Sensor Syst.* New York, NY, USA: Association for Computing Machinery, Nov. 2017, pp. 1–2, doi: 10.1145/3131672.3136993.
[62] R. Bhatt. (2017). *Wireless Indoor Localization*. UCI Machine Learning Repository. [Online]. Available: https://doi.org/10.24432/C51880
[63] J. G. Rohra, B. Perumal, S. J. Narayanan, P. Thakur, and R. B. Bhatt, ''User localization in an indoor environment using fuzzy hybrid of particle swarm optimization & gravitational search algorithm with neural networks,'' in *Proc. 6th Int. Conf. Soft Comput. Problem Solving, SocProS*, vol. 1. Cham, Switzerland: Springer, 2017, pp. 286–295.

XINYUE WANG received the degree from the Shandong University of Science and Technology. She is currently a communication engineering research scholar. Her research interest includes use machine learning methods to build communication and positioning systems for the Internet of Things. She has participated in a number of internal scientific research and competition projects and accumulated a certain amount of research experience. She have high practical and research ability in this field. In her research, a large number of machine learning algorithms, including deep learning and convolutional neural networks, are used to build the IoT communication and positioning systems. These systems are designed to increase the speed and correctness of data transmission and improve the reliability and accuracy of communication and positioning systems. She is also a passionate and innovative scientist who is committed to advancing the development of IoT technology and has made active contributions in the field of communication engineering.