3-D BLE Indoor Localization Based on Denoising Autoencoder

CHAO XIAO[1], DAIQIN YANG[1], (Member, IEEE), ZHENZHONG CHEN[1], (Senior Member, IEEE), AND GUANG TAN[2], (Member, IEEE)

[1]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

[2]SIAT, Chinese Academy of Sciences, Shenzhen 518005, China

Corresponding author: Daiqin Yang (dqyang@whu.edu.cn)

ABSTRACT Bluetooth low energy (BLE)-based indoor localization has attracted increasing interests for its low-cost, low-power consumption, and ubiquitous availability in mobile devices. In this paper, a novel denoising autoencoder-based BLE indoor localization (DABIL) method is proposed to provide high-performance 3-D positioning in large indoor places. A deep learning model, called denoising autoencoder, is adopted to extract robust fingerprint patterns from received signal strength indicator measurements, and a fingerprint database is constructed with reference locations in 3-D space, rather than traditional 2-D plane. Field experiments show that 3-D space fingerprinting can effectively increase positioning accuracy, and DABIL performs the best in terms of both horizontal accuracy and vertical accuracy, comparing with a traditional fingerprinting method and a deep learning-based method. Moreover, it can achieve stable performance with incomplete beacon measurements due to unpredictable BLE beacon lost.

INDEX TERMS Bluetooth low energy, indoor localization, fingerprinting, denoising autoencoder.

I. INTRODUCTION

With the rapid development of mobile Internet, location based services (LBS) in large public indoor places, such as shopping malls, offices, and airports, have become increasingly popular. To date, a lot of indoor localization technologies have been proposed based on various measuring techniques, including ultrasound [1], [2], infrared [3], image [4], [5], light [6], magnetic field [7], and wireless signals [8]–[16]. For example, Bordoy *et al.* [1] presented an indoor localization approach based on measuring the time of flight (TOF) of ultrasound signal reflections in human bodies and walls. Similarly, ultrasonic signals can also be used to estimate distance between receiver and transmitter by measuring the TOF [2]. Gorostiza *et al.* [3] presented a sensorial system to perform indoor localization by measuring differential phase-shifts of sinusoidally modulated infrared signals from a robot. Zheng *et al.* [4] developed a system that uses RGB-D cameras for visual based localizations. Werner *et al.* [5] proposed a positioning method which allows for a fine-grained detection of the position and orientation of users with smartphone cameras. In SpinLight [6], spatial beams are identified with a unique time sequence of light signals generated by a coded shade which cover and rotate around an infrared LED light source. LocateMe [7] uses magnetic field as spatial patterns

for indoor localization. In addition to the diversified measuring primitives above, wireless signals, especially WiFi signal, are very popular in indoor localization due to their wide availability on mobile devices and indoor environments. In [8], EZ models wireless propagation with the log-distance path loss (LPDL) model and uses a genetic algorithm for indoor localization. Li *et al.* [9] proposed a WiFi indoor localization method based on collaboration of fingerprinting and assistant nodes. The authors in [10] proposed a WiFi localization system based on the channel state information (CSI) of WiFi signals and the deep belief network (DBN). The deep learning model is used to learn the fingerprint patterns from high dimensional CSI signals. Shu *et al.* [11] proposed gradient fingerprinting (GIFT) which leverages a more stable RSSI gradient. Liu *et al.* [12] proposed a peer assisted localization approach to eliminate large errors of WiFi signals. Beside WiFi, other wireless signals are also widely used by indoor localization technologies, including ZigBee, UWB, and RFID. For example, Sugano *et al.* [13] implemented an autonomous indoor localization system for wireless sensor network with ZigBee signals. In [14], a method was proposed to estimate the Time-Difference-of-Arrival (TDOA) for UWB indoor positioning. Wang *et al.* [15] proposed a robust algorithm named residual weighted multidimensional

scaling (RWMDS) based on multipath channel model for passive UHF RFID tag localization. Saab and Nakad [16] also used RFID tags to implement indoor positioning with Kalman filter.

Till now, most of the indoor localization methods are designed for 2D localizations only. In some special applications, such as indoor navigation of unmanned aerial vehicle (UAV), commodities positioning and tracking in large supermarkets, and so on, 3D indoor localization is very important. To solve this problem, several technologies have been proposed with different measuring primitives [17]–[19]. For example, Lu et al. [17] introduced a 3D Structure-from-Motion (SfM) model for 3D indoor localization from videos records. Xu et al. [18] designed and implemented a RFID based 3D indoor localization system using RSSI measurements for multiple trilateral positioning. In [19], a visible light based 3D indoor positioning system was proposed, adopting code division multiple access (CDMA) technique for spatial positioning from multiple reference light sources.

Recently, Bluetooth Low Energy based localization has attracted intense interests and gained popularity in practical use due to its low cost, low power consumption, and ubiquitous availability in mobile devices. Similar with other wireless techniques, BLE can be used from two different approaches: trilateration [20] and fingerprinting [21], [22]. The *trilateration approach* exploits the relationship between the receiver's RSSI measurement and its distance to the BLE beacon station. By estimating distances to multiple beacon stations, the receiver's location can be recovered with a least square algorithm. The *fingerprinting approach* is implemented in two phases: an offline phase and an online phase. The offline phase is called the training phase. During this phase, a fingerprint database is constructed with position-dependent parameters extracted from the measured RSSIs of reference locations. The online phase is known as the localization phase. It maps the current RSSI measurements to a reference location by finding the most relevant RSSI fingerprint from the database. However, as wireless signals are not stable and may vary greatly across location and time, both of these two approaches suffer from large positioning errors. Moreover, BLE suffers severe beacon lost comparing with WiFi and other wireless techniques due to the following three reasons. Firstly, there are only three broadcast channels allocated for BLE beacons. When there are multiple beacon stations installed in vicinity of the receiver, collisions may easily happen such that the receiver cannot recover the beacon signals correctly. Secondly, with much lower transmission power of BLE beacons, the transmission range of BLE is much shorter than that of WiFi. It's thus more likely for receivers to be in the boundary areas of BLE beacon stations and suffer from beacon lost due to fluctuations of receiving signal power. Thirdly, BLE beacon stations are usually battery powered, so there would be signal vacuums when the batteries run out.

To address these problems, a novel denoising autoencoder based BLE indoor localization method, DABIL, is proposed in this paper. A deep learning model, called denoising autoencoder [23], is adopted to extract robust fingerprint patterns from noisy RSSI measurements. Based on these robust fingerprints, a 3D fingerprint database is constructed offline for reference locations in 3D space, to provide accurate 3D online positioning. As demonstrated by our experimental results, a 3D fingerprinting database can effectively improve localization accuracies, in terms of both horizontal and vertical accuracy. Extensive performance evaluations show that DABIL outperforms other methods in both horizontal and vertical location accuracies.

The remainder of this paper is organized as follows. Section II presents technical details of DABIL. Experiment results are demonstrated and discussed in Section III. Section IV concludes our work.

II. DEEP LEARNING BASED BLE INDOOR LOCALIZATION
DABIL adopts the fingerprinting approach. Comparing with trilateration approach, fingerprinting exhibits robust performances against multi-path effect of indoor radio propagation. Complex layout of indoor environment, including walls, furniture, decorations, etc., can bring complex reflections to indoor radio propagations, and multiple copies of the same signal may reach the receiver side from different reflection paths, each with different attenuations and delays. The unpredictable weakening or strengthening combination effect of these multi-path signals will break the relationship between RSSI and transmission distance, and thus render the trilateration positioning less effective. Fingerprinting approach does not rely on the recovery of distances. It uses the measured RSSIs as spatial patterns only and is thus less vulnerable to the multi-path effect.

A. FINGERPRINTING
The key step in the fingerprinting method is to build a fingerprint database. A typical fingerprint database can be expressed as:

$$\Omega = \{(\mathbf{f}_1, \mathbf{p}_1), (\mathbf{f}_2, \mathbf{p}_2), \ldots, (\mathbf{f}_N, \mathbf{p}_N)\} \quad (1)$$

where N is the total number of reference locations in the database, $\mathbf{f}_i$ denotes the fingerprint pattern of the ith reference location and $\mathbf{p}_i$ is the spatial coordinates of this reference location.

The fingerprint pattern $\mathbf{f}$ can be raw RSSI values from multiple beacon stations, or any other feature vectors extracted from those RSSIs. Most of the existing fingerprinting systems use raw RSSI values as spatial patterns. In this case, $\mathbf{f}$ can be expressed as:

$$\mathbf{f} = [r_1, r_2, \ldots, r_m] \quad (2)$$

where m is the total number of BLE beacon stations, and r_i represents the measured RSSI value from the ith station. This fingerprinting is simple but suffers from several drawbacks. First, the set of covering beacon stations could vary violently across locations and it is hard to divide them into different subsets in the RSSI space. Second, the value of
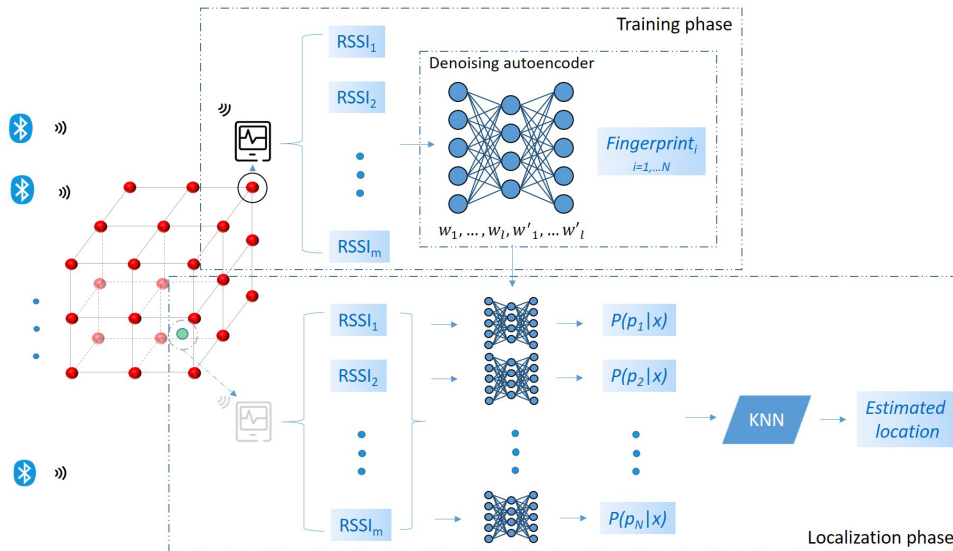
FIGURE 1. System architecture of DABIL.

RSSIs fluctuate frequently and widely. In DABIL, a denoising autoencoder is used in the training phase, which can learn features from measured RSSIs with its simple and symmetrical structure and behave robustly against signal fluctuation as well as beacon lost.

In other words, a deep network is trained with fluctuated raw RSSI measurements, and the trained network itself is used as the fingerprint pattern for this reference location. Since a deep network can be represented by its weights of each layer, the fingerprint pattern of DABIL can be expressed as follows:

$$\mathbf{f} = [w_1, w_2, \ldots, w_l, w'_1, w'_2, \ldots, w'_l] \tag{3}$$

where l is the number of encoding hidden layers of the denoising autoencoder, w_i and w'_i represent the weights of the ith encoding hidden layer and its decoding mirror layer.

The architecture of DABIL is shown in Fig. 1. Different from other methods, RSSIs of BLE beacon stations are measured at different heights in the training phase. At each reference location, a serial of RSSI measurements are first collected from all the nearby BLE beacon stations and then used to train a denoising autoencoder for fingerprint. In the localization phase, these trained denoising autoencoders are used as fingerprints to calculate the probabilities that the target is located at each reference location. And finally, the estimated target location is calculated by a KNN algorithm.

B. TRAINING OF FINGERPRINT USING DENOISING AUTOENCODER

Deep learning [24], also known as deep structured learning, is a branch of machine learning based on a set of algorithms that attempt to model high-level abstractions in data by using a deep graph with multiple processing layers, composed of multiple linear and non-linear transformations. It has been widely used to learn features from non-structed data, such as images [25], sounds [26], EEG signals [27], and texts [28]. An autoencoder [29] is an artificial neural network designed to learn efficient encodings of an original data set. And it is widely used for the purpose of dimensionality reduction and feature extraction.

The autoencoder is an unsupervised learning method, the input and output layers of which are of the same size. With encoding hidden layers, it predicts a target $\mathbf{z}$ with given input $\mathbf{x}$, and then reconstructs the input with $\mathbf{z}$ through decoding mirror layers. By defining the encoding and decoding functions as $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$, the process can be expressed as:

$$\mathbf{z} = \boldsymbol{\phi}(\mathbf{x}) \tag{4}$$
$$\mathbf{x}' = \boldsymbol{\psi}(\mathbf{z}) \tag{5}$$

where $\mathbf{x}'$ is the reconstructed input. The aim of an autoencoder is to minimize the difference between the original and reconstructed input.

$$\min_{\boldsymbol{\phi}, \boldsymbol{\psi}} \|\mathbf{x} - \boldsymbol{\psi}(\boldsymbol{\phi}(\mathbf{x}))\|^2$$

In DABIL, input $\mathbf{x}$ is a column vector of measured RSSIs from multiple beacon stations at a reference location:

$$\mathbf{x} = [r_1, r_2, \ldots, r_m]^T \tag{6}$$

where r_i is the RSSI value measured from the ith BLE beacon station and m is the number of BLE beacon stations.

Denoising autoencoder [23] is a stochastic version of autoencoder. It tries to encode the input in a way that can undo the effect of a stochastic corruption process applied to it. This requires that the trained network can capture the statistical dependencies among different inputs. It has two obvious advantages. Firstly, it can cast away the dependency among input dimensions for better feature extraction. Secondly, it is
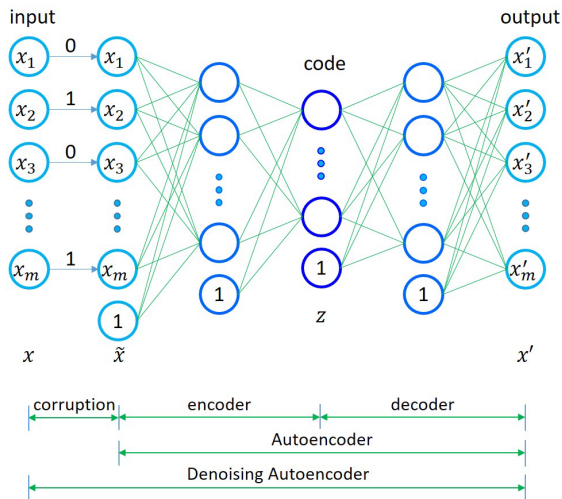
FIGURE 2. Structure diagram of a denoising autoencoder.

more stable when there are noises in the input. For these reasons, the denoising autoencoder is adopted in DABIL to both improve localization accuracy and provide robustness with RSSI fluctuations and beacon lost.

To convert an autoencoder into a denoising autoencoder, a stochastic corruption step should be added to the training input, which is shown in Fig. 2. The training input can be corrupted in many ways. In DABIL, the input is corrupted in the way that entries of $\mathbf{x}$ are randomly set to zero. This corruption function is designed to simulate the randomized loss of BLE beacon signals.

During the training phase, a denoising autoencoder with three encoding hidden layers is constructed for each reference location and trained with multiple measurements. All these trained denoising autoencoders are used as fingerprints to construct the fingerprint database. A denoising autoencoder with three encoding hidden layers cannot be trained in one step. It should be trained in a layer-by-layer fashion. Let z_1, z_2, z_3 denote the output of the first, second and third encoding hidden layer, as shown in Fig. 3. The training process can be described as the following steps:

0) Corrupt the input.
   Corrupt the input $\mathbf{x}$ to $\tilde{\mathbf{x}}$ with:

$$\tilde{\mathbf{x}} = \mathbf{w}_0 \odot \mathbf{x} \tag{7}$$

where $\odot$ represents entry-wise production, and $\mathbf{w}_0$ is a column vector whose entries are randomly set to either 0 or 1 with a probability of P_c and $1-P_c$. In the following training process, P_c is set to 50%. It is noteworthy that $\mathbf{w}_0$ is regenerated in each training batch, insuring stochastic corruptions so as to reduce the dependencies between every two entries of the input as much as possible.

1) Train the first hidden layer.
   Construct a three-layer artificial neural network with the corrupted input layer, the first encoding hidden layer,

and its decoding mirror layer. The corrupted input $\tilde{\mathbf{x}}$ is first mapped to $\mathbf{z}_1$ with:

$$\mathbf{z}_1 = \sigma(\mathbf{w}_1 \tilde{\mathbf{x}} + \mathbf{b}_1) \tag{8}$$

where $\mathbf{w}_1$ is the weighting matrix of the first encoding hidden layer, and $\mathbf{b}_1$ is its bias vector. σ is the activation function of each node in the hidden layer, which is sigmoid function in DABIL. After this, $\mathbf{z}_1$ is further mapped to reconstruct $\mathbf{x}$. Denoting the reconstructed input as $\mathbf{x}'$, we can have

$$\mathbf{x}' = \sigma(\mathbf{w}_1' \mathbf{z}_1 + \mathbf{b}_1') \tag{9}$$

where $\mathbf{w}_1'$ is the weighting matrix of the decoding mirror layer, and $\mathbf{b}_1'$ is its bias vector. This three layer network is trained to minimize the reconstruction error of:

$$\mathcal{L}(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|^2$$
$$= \|\mathbf{x} - \sigma(\mathbf{w}_1'(\sigma(\mathbf{w}_1 \tilde{\mathbf{x}} + \mathbf{b}_1)) + \mathbf{b}_1')\|^2 \tag{10}$$

The gradient descent algorithm is used for the adjustment of weights, and the network is trained for 100 times or until the error is less than 0.001.

2) Train the second hidden layer.
   Construct a three-layer artificial neural network with input $\mathbf{z}_1$, the second encoding hidden layer, and its decoding mirror layer. Similarly, the input $\mathbf{z}_1$ is mapped to $\mathbf{z}_2$ and $\mathbf{z}_2$ is mapped to a reconstructed $\mathbf{z}_1'$:

$$\mathbf{z}_2 = \sigma(\mathbf{w}_2 \mathbf{z}_1 + \mathbf{b}_2) \tag{11}$$
$$\mathbf{z}_1' = \sigma(\mathbf{w}_2' \mathbf{z}_2 + \mathbf{b}_2') \tag{12}$$

and the network is trained to minimize the reconstruction error of:

$$\mathcal{L}(\mathbf{z}_1, \mathbf{z}_1') = \|\mathbf{z}_1 - \mathbf{z}_1'\|^2$$
$$= \|\mathbf{z}_1 - \sigma(\mathbf{w}_2'(\sigma(\mathbf{w}_2 \mathbf{z}_1 + \mathbf{b}_2)) + \mathbf{b}_2')\|^2 \tag{13}$$

3) Train the third hidden layer.
   The third hidden layer of the denoising autoencoder is trained in the same way as Step 2, with the minimization target of:

$$\mathcal{L}(\mathbf{z}_2, \mathbf{z}_2') = \|\mathbf{z}_2 - \mathbf{z}_2'\|^2$$
$$= \|\mathbf{z}_2 - \sigma(\mathbf{w}_3'(\sigma(\mathbf{w}_3 \mathbf{z}_2 + \mathbf{b}_3)) + \mathbf{b}_3')\|^2 \tag{14}$$

After the whole denoising autoencoder is trained, the weighting matrix and bias vector of each layer is obtained, and the fingerprint pattern of the ith reference location can be represented as:

$$\mathbf{f}_i = [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_1', \mathbf{w}_2', \mathbf{w}_3']. \tag{15}$$

where the bias vectors are combined into each weighting matrix as an extra row. Then the fingerprint database is constructed with the fingerprint patterns and position coordinates of all the reference locations.
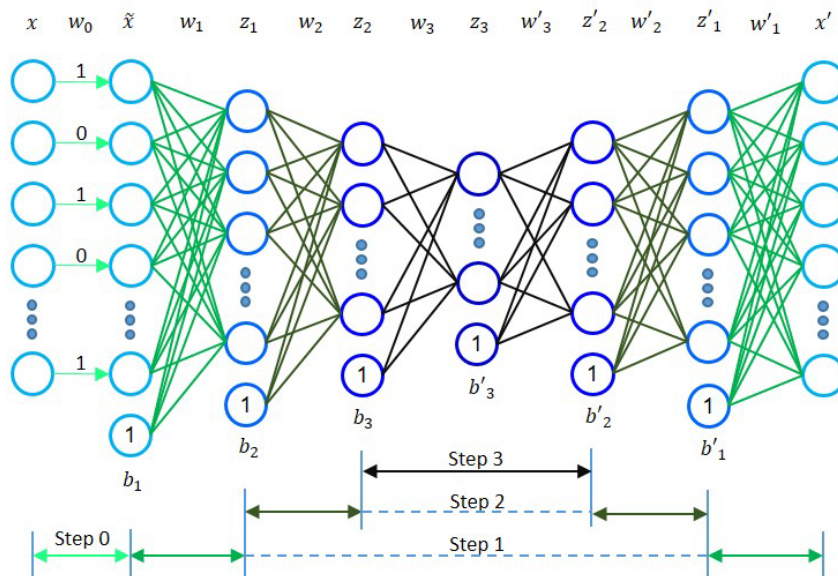
FIGURE 3. Training process of the denoising autoencoder.

C. LOCATION ESTIMATION

To estimate a target location, the posteriori probabilities that the target locates at each reference locations can be calculated first. Then the target location can be estimated with a weighted KNN method as follows:

$$\hat{\mathbf{p}} = \sum_{i=1}^{k} P(i|\mathbf{x})\mathbf{p}_i \qquad (16)$$

where $P(i|\mathbf{x})$ is the posteriori probability of target locating at reference location i, and $\mathbf{p}_i$ is the coordinate of the ith reference location. In existing fingerprinting methods, the heights of fingerprints are always ignored and the fingerprints are usually collected at a fixed height, i.e., $\mathbf{p}_i = [x_i, y_i]$. In fact, RSSI varies across different heights. Without taking height information into consideration, the positioning accuracy could fluctuate violently across different heights. So in our method, a 3D fingerprinting approach is adopted to deal with this issue. Fingerprints are collected in 3D space and each reference location has its 3D coordinates $\mathbf{p}_i = [x_i, y_i, z_i]$. This design can both provide height information and improve horizontal localization accuracy.

According to Bayes' law, we can get:

$$P(i|\mathbf{x}) = \frac{P(i)P(\mathbf{x}|i)}{\sum_i P(i)P(\mathbf{x}|i)} \qquad (17)$$

Here, $P(i)$ is the prior probability that the target point is located at the ith reference location. Assuming that $P(i)$ is uniformly distributed for each reference location, (17) can be simplified as:

$$P(i|\mathbf{x}) = \frac{P(\mathbf{x}|i)}{\sum_i P(\mathbf{x}|i)} \qquad (18)$$

For each reference location i, a denoising autoencoder is built with its fingerprint pattern $\mathbf{f}_i$, and the measured RSSIs at the target location are used as the input $\mathbf{x}$ of the denoising autoencoder. In this phase, the corruption step is not required anymore. And the distance is calculated between the input $\mathbf{x}$ and its reconstructed version of $\mathbf{x}'_i$:

$$d_i = \|\mathbf{x} - \mathbf{x}'_i\| \qquad (19)$$

There are relationships between $P(\mathbf{x}|i)$ and d_i. Fig. 4 shows the mean value and standard deviation of d_i with different D_i, which is the 3D space distance between the target location and the ith reference location. From Fig. 4, it is found that d_i increases generally with D_i, though there are some variations due to the instability of BLE signals. It can thus be inferred that the larger d_i is, the larger D_i is, and consequently the smaller $P(\mathbf{x}|i)$ is. $P(\mathbf{x}|i)$ is therefore defined as,

$$P(\mathbf{x}|i) = \exp\left(-\frac{d_i}{\lambda}\right) \qquad (20)$$

the curve of which is shown in Fig. 5. The choice of this exponential function can be roughly attributed to two reasons. Firstly, $P(\mathbf{x}|i)$ decreases with d_i under this function, which conforms to the findings from Fig. 4. Secondly, $P(\mathbf{x}|i)$ falls into the range of $(0, 1]$ for positive d_i, and the shape of the function conforms to our intuitive of the probability distributions. In Fig. 4, the values of d_i are within the range of $(0.1, 0.7)$ when D_i is in the range of $(0, 8)$. Taking into consideration that the coverage range of BLE is around 10 meters, and reference locations outside coverage range are of no use as their signals cannot be correctly recognized. Combining with Fig. 5, it can be found that the shape of the curve gives a better discrimination for the range of $d_i \in (0.1, 0.7)$ with $\lambda = 1/3$. The distribution range of $P(\mathbf{x}|i)$ is too small when $\lambda = 1/2$, and the shape of the function is too steep when $\lambda = 1/4$. So in this paper, the value of λ is set to $1/3$.
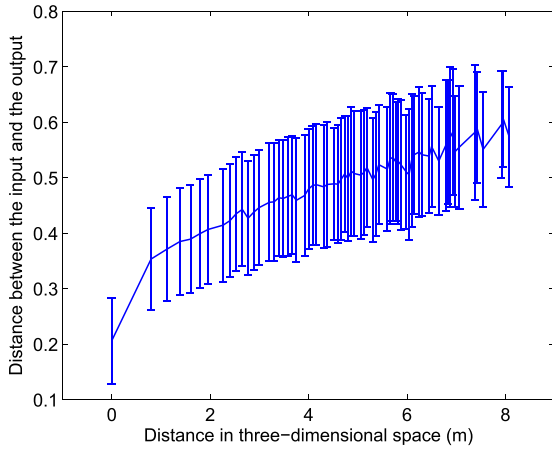
FIGURE 4. Coding distance *d* vs location distance *D*.
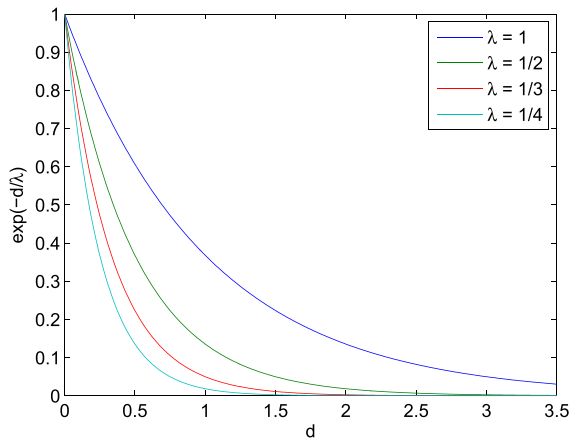

FIGURE 5. The curve of $\exp\left(-\frac{d}{\lambda}\right)$.

With the calculated value of each $P(\mathbf{x}|i)$, the estimated position can be calculated with Eq. (16), which is represented as following:

$$\hat{\mathbf{p}} = \sum_{i=1}^{k} \frac{P(\mathbf{x}|i)}{\sum_i P(\mathbf{x}|i)} \mathbf{p}_i \qquad (21)$$

III. PERFORMANCE EVALUATION

As shown in Fig. 6, a testbed of BLE beacon system are deployed in a conference room for the purpose of performance evaluation. The layout of the conference room measures 17.5×9.6 meters. Ten BLE beacon stations are placed on the wall at height of 2.54 meters. Reference locations (red points in Fig. 6) are measured at three different heights of 0.8m, 1.6m, and 2.4m, respectively. The distance between adjacent reference locations from a same height is 0.8m. Thus, the total number of the reference points is $8 \times 8 \times 3 = 192$. Two layers of test locations (green points in Fig. 6) are measured at heights of 1.2m and 1.9m respectively.

A. RSSI FILTERING AND PARAMETER SELECTION

The raw RSSI readings are not directly used due to its severe fluctuation across time. Fig. 7 shows the measured RSSIs of a same BLE beacon station at different time. It shows that

RSSIs are unstable and have strong signal fluctuations due to fading and multi-path effects. Besides this, there could also be beacon lost during the measurement of RSSIs. So, in DABIL, a filter with time window T is used to smooth the raw RSSI readings and eliminate the effect of RSSI fluctuation and beacon lost. The filter can be mean filter, median filter, or other low pass filters. In DABIL, a mean filter is adopted with $T = 1s$.

To achieve high localization accuracy and robustness, it's essential to choose the optimal model parameters. The input node number of the denoising autoencoder is set to ten, which is the same as the installed beacon stations. And there are four parameters left in DABIL to be determined, including the node numbers of three different hidden layers in the denoising autoencoder and the number of neighbours k in KNN. Let n_1, n_2, n_3 be the node numbers of the first, second and third hidden layers respectively. Taking into consideration that a denoising autoencoder can be seen as a process of feature extraction, the node number of its hidden layers usually decrease along the data pass. The range of parameters are thus roughly set as $n_1 \in [8, 18]$, $n_2 \in [5, 14]$, $n_3 \in [3, 10]$, and $k \in [1, 30]$ for an affordable exhaustion. Using the localization accuracy as selection criterion, the optimal configuration is found to be $n_1 = 9$, $n_2 = 13$, $n_3 = 5$, and $k = 9$. Fig. 8 and Fig. 9 depict the effect of value k to the localization accuracy of DABIL, in terms of mean and standard deviation of localization errors. As shown in these two figures, $k = 9$ provides a balance among precision, stability, and computational complexity.

B. LOCALIZATION ACCURACY

The performance of DABIL is compared with Faragher's method [22] and Zhang's method [30]. Faragher's method is a widely accepted BLE fingerprinting method, which was designed with reference locations in 2D plane (RL-2D) and supports 2D localization only. DABIL uses reference locations in 3D space (RL-3D) and provides 3D localization. In Zhang's method, a four-layer Deep Neural Network (DNN) and a Hidden Markov Model (HMM) are combined to realize indoor localization. The DNN is used as a classifier to achieve coarse localization, which can be deemed as fingerprinting, and the HMM is introduced to smooth the initial localization so as to improve localization accuracy. For the purpose of a fair performance evaluation, Faragher's method and Zhang's method are tested with RL-3D setting, and DABIL is also tested with RL-2D setting. Denoting $[x_e, y_e, z_e]$ and $[x_r, y_r, z_r]$ as the estimated and real coordinates of the test location, localization accuracies are compared in terms of horizontal error $E_h = \sqrt{((x_e - x_r)^2 + (y_e - y_r)^2)}$ and vertical error $E_v = |z_e - z_r|$ respectively, so as to accommodate 2D and 3D localization into a same comparison. Besides these, a degenerated version of DABIL, which is trained with basic autoencoder only, is also tested to verify the robustness brought by denoising autoencoder.
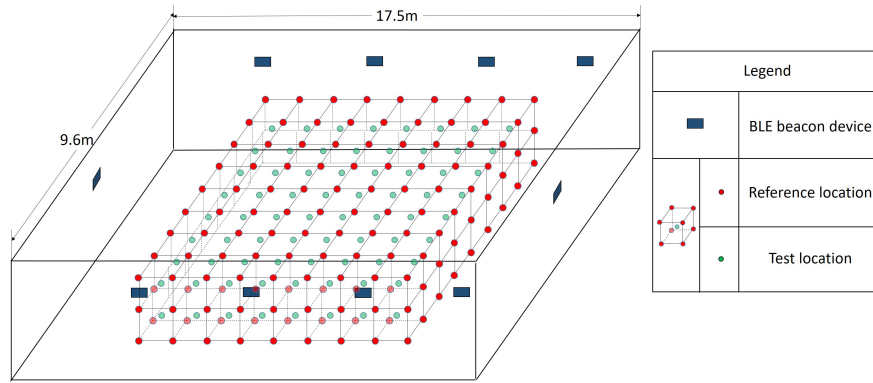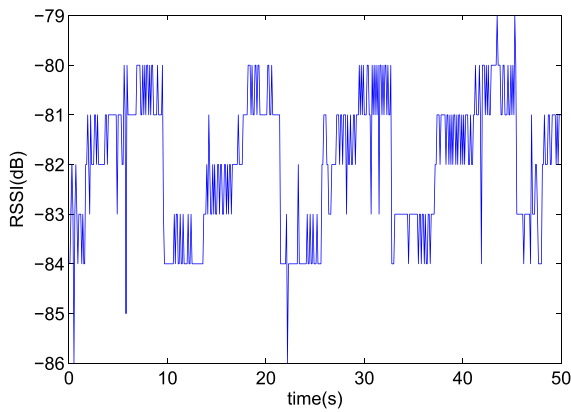
FIGURE 6. Testbed layout.
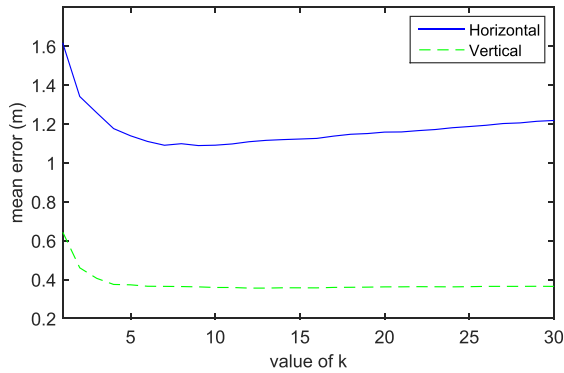

FIGURE 7. Time variation of BLE RSSIs.

FIGURE 8. Mean localization error vs k.

FIGURE 9. Standard deviation of localization error vs k.

FIGURE 10. CDFs of horizontal localization accuracy.

1) HORIZONTAL LOCALIZATION ACCURACY

As shown in Fig. 10 and Table 1, DABIL performs the best in 2D positioning. It achieves distance errors of 1.0m and 2.0m for 52.69% and 92.56% of total instances, and its mean accuracy is about 1.09m. Faragher's method produces location errors of 1.0m and 2.0m for only 32.82% and 76.28% of the instances. The mean accuracy is about 1.51m, which is about 50% worse than that of DABIL. Zhang's method produces location errors of 1.0m and 2.0m for 48.69% and 89.76% of the instances, and the mean accuracy is 1.18m, which is slightly inferior to that of DABIL. When trained
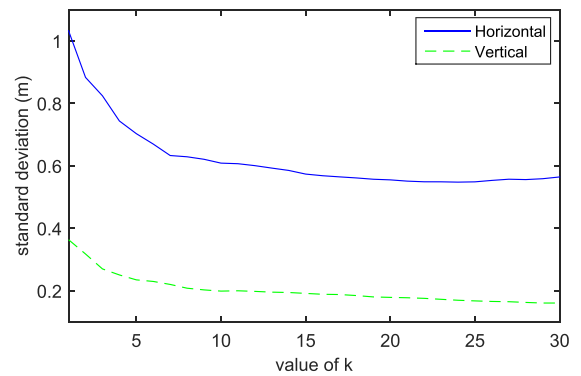
with autoencoder only, the positioning accuracy of DABIL decreases to 48.72% and 86.15% for 1.0m and 2.0m accuracy, and the mean accuracy decreases to about 1.23m. The performance gap between denoising autoencoder and autoencoder shows the improved stability by using denoising autoencoder for fingerprint training.

When applying RL-3D settings to Faragher's method, it could be found that the location accuracy improves

TABLE 1. Horizontal mean accuracy and standard deviation.

	Mean Error	Std. Dev.
DABIL	1.0891	0.6211
DABIL with Autoencoder	1.2292	0.6957
DABIL with RL-2D	1.2980	0.6707
Faragher's Method	1.5104	0.8574
Faragher's Method with RL-3D	1.2262	0.7685
Zhang's Method	1.1814	0.9169
Zhang's Method with RL-3D	1.2462	0.8430

significantly comparing with its RL-2D settings, though not as good as DABIL. 47.82% and 86.03% of the total instances achieve positioning errors within 1.0m and 2.0m, and the mean accuracy is about 1.23m. These experimental results demonstrate that by collecting fingerprint in 3D space, the accuracy of 2D positioning with traditional fingerprinting can be increased. When applying RL-2D settings to DABIL, the location accuracy also decreases, though still better than that of Faragher's method. 38.59% and 83.97% of the total instances achieve position errors of 1.0m and 2.0m, and the mean accuracy is about 1.30m. When applying RL-3D settings to Zhang's method, the accuracy decreases to 45.36% and 85.95% for 1.0m and 2.0m, and the mean accuracy decreases to about 1.25m. The reason for this unexpected degradation of localization accuracies can be attributed to the HMM used by the algorithm. With RL-3D settings, the number of states in the HMM would be tripled comparing with RL-2D settings. This enlarged state number, together with the error propagation effect of HMM, aggravate the accuracy of Zhang's method.

TABLE 2. Vertical mean accuracy and standard deviation.

	Mean Error	Std. Dev.
DABIL	0.3410	0.2027
DABIL with Autoencoder	0.3626	0.2029
Faragher's Method with RL-3D	0.3726	0.2419
Zhang's Method with RL-3D	0.5000	0.3224

2) VERTICAL LOCALIZATION ACCURACY

To compare performances in terms of vertical accuracy, Faragher's method and Zhang's method are modified to provide 3D positioning with RL-3D settings. Fig. 11 and Table 2 show the comparison results. DABIL still performs the best with 36.41% and 92.44% for 0.2m and 0.6m accuracy, and its mean accuracy is about 0.34m. Faragher's method performs slightly worse, with 36.79% and 84.74% for 0.2m and 0.6m height errors. Its mean vertical accuracy is about 0.37m. When trained with basic autoencoder, DABIL's vertical accuracy slightly decreases to the level of 33.97% and 90.26% for 0.2m and 0.6m. And the mean accuracy is about 0.36m, which is still better than that of Faragher's. Zhang's method performs the worst as expected, with 11.67% and 76.43% for 0.2m and 0.6m height errors, and its mean accuracy is 0.50m, which further proves that HMM is not compatible with RL-3D settings.
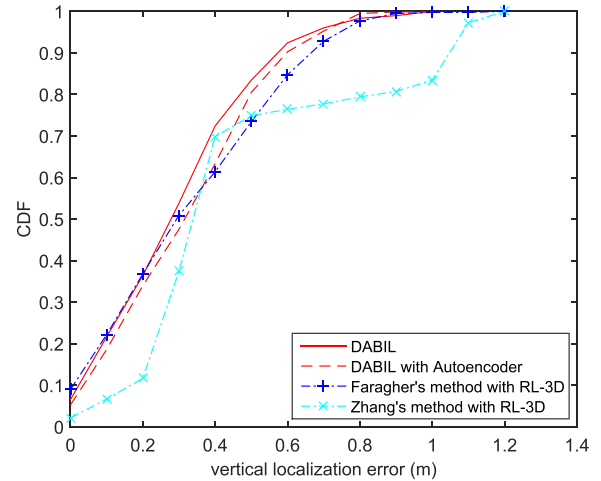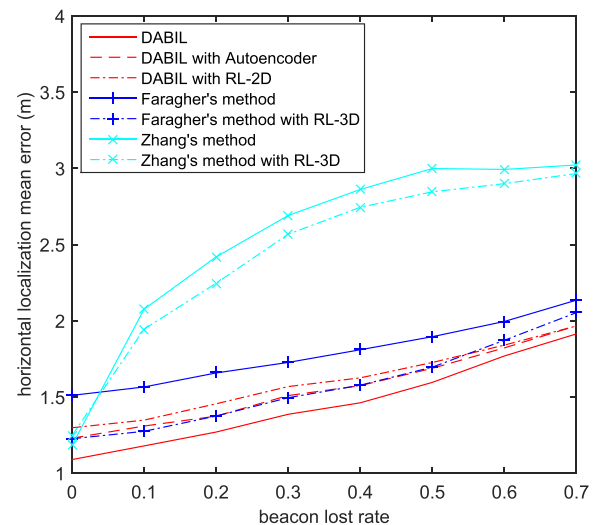

FIGURE 11. CDFs of vertical localization accuracy.

FIGURE 12. Horizontal localization accuracy vs beacon lost rate.

3) LOCALIZATION ACCURACY IN CASE OF BEACON LOST

To validate the robustness of DABIL in case of beacon lost, entries of measured RSSIs are stochastically masked by zeros with a certain probability (beacon lost rate). Corrupted inputs are used to test the robustness of DABIL and other algorithms. The results are shown in Fig. 12 and Fig. 13. These emulated results demonstrate that DABIL performs the best among all the evaluated methods despite the beacon lost rate.

In terms of horizontal localization, DABIL can retain accuracy level of 1.27m and 1.60m when the beacon lost rates are 20% and 50% respectively. When trained with basic autoencoder only, the positioning accuracy of DABIL decreases to about 1.37m and 1.69m for 20% and 50% beacon lost. Faragher's method performs even worse. The accuracy is only about 1.66m and 1.90m. With RL-3D settings, the accuracy of Faragher's method increases to 1.37m and 1.70m, which are still worse than that of DABIL. Zhang's method is very vulnerable to the corruption of inputs and its accuracy acutely
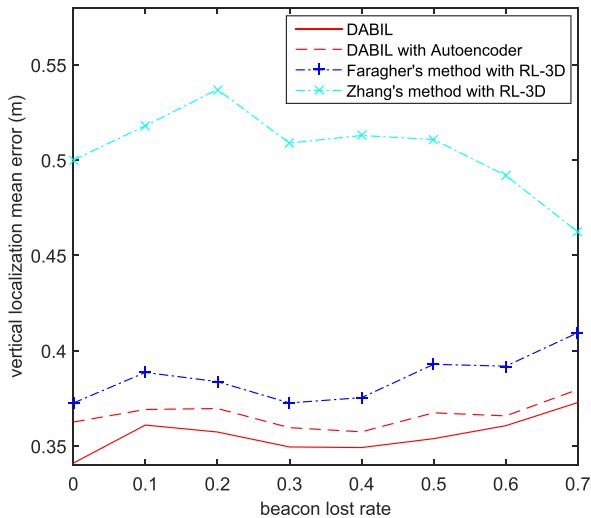
FIGURE 13. Vertical localization accuracy vs beacon lost rate.

TABLE 3. Computational time per instant.

	Mean Time (ms)
DABIL	0.84
DABIL with Autoencoder	0.81
DABIL with RL-2D	0.28
Faragher's Method	0.06
Faragher's Method with RL-3D	0.08
Zhang's Method	1.1
Zhang's Method with RL-3D	3.8

decreases to 2.42m and 3.00m for 20% and 50% beacon lost rate. With RL-3D settings, Zhang's method is not stable and its accuracy is about 2.25m and 2.85m for 20% and 50% beacon lost rate.

For vertical performance, it is found that the localization error does not increase monotonously with beacon lost rate. This can be attributed to the relatively small range of reference locations in the vertical direction. Nevertheless, the general tendency of localization error is increasing as expected, excluding Zhang's method, and DABIL still performs the best under all beacon lost rates.

C. SPEED ANALYSIS
Computational complexities of each algorithm are compared in terms of time needed for the calculation of each location. Average results of 840 instances are listed in Table 3. As shown in Table 3, Faragher's method costs the least computational time as expected due to its simple localization process, i.e. 0.08ms and 0.06ms for RL-3D and RL-2D settings. DABIL and Zhang's method demand much longer calculation time, even though their major computational burdens are in the offline training phase and not counted in the results. DABIL with denoising autoencoder can estimate the target's location in 0.84ms, and DABIL with autoencoder costs 0.81ms. This sub-second calculation time can meet the requirement of real-time indoor localization. When use RL-2D settings only, DABIL only needs 0.28ms to calculate one position, which is about 1/3 that of using RL-3D settings. Zhang's method costs the longest calculation

time, with 3.8ms and 1.1ms for RL-3D and RL-2D settings. Generally, with much less reference locations, computational time with RL-2D settings are shorter than that of RL-3D settings.

IV. CONCLUSION
In this paper, a novel denoising autoencoder based BLE indoor localization method, DABIL, is proposed. Reference locations are collected in 3D space instead of commonly used 2D plane. A deep learning model, denoising autoencoder, is adopted for robust fingerprint extraction from measured RSSIs. Denoising autoencoders are trained with RSSIs of reference locations, and the weights of each trained denoising autoencoder are used as the fingerprint of this reference location. By sending the measured RSSIs of a target location into the trained denoising autoencoders of each reference location, the distances between the outputs and the input can be used indicators of the likelihood of each reference location. A KNN method is then applied for location estimation with a weighted average of those most related reference locations. Field experiments are conducted to evaluate the accuracy of DABIL and its robustness against beacon lost. Experimental results demonstrate that, DABIL outperforms Faragher's method [22] and Zhang's method [30] in both location accuracies (horizontal and vertical) and robustness against beacon lost. Besides, 3D reference locations can help increase traditional fingerprinting localization accuracy effectively, comparing with 2D settings.

REFERENCES
[1] J. Bordoy, J. Wendeberg, C. Schindelhauer, and L. M. Reindl, "Single transceiver device-free indoor localization using ultrasound body reflections and walls," in *Proc. IEEE IPIN*, Oct. 2015, pp. 1–7.

[2] C. Medina, J. C. Segura, and Á. De la Torre, "Ultrasound indoor positioning system based on a low-power wireless sensor network providing sub-centimeter accuracy," *Sensors*, vol. 13, no. 3, pp. 3501–3526, 2013.

[3] E. M. Gorostiza, J. L. L. Galilea, F. J. M. Meca, D. S. Monzú, F. E. Zapata, and L. P. Puerto, "Infrared sensor system for mobile-robot positioning in intelligent spaces," *Sensors*, vol. 11, no. 5, pp. 5416–5438, 2011.

[4] Y. Zheng, P. Luo, S. Chen, J. Hao, and H. Cheng, "Visual search based indoor localization in low light via RGB-D camera," *Int. J. Comput. Electr. Autom. Contr. Inf. Eng.*, vol. 11, no. 3, pp. 349–352, 2017.

[5] M. Werner, M. Kessel, and C. Marouane, "Indoor positioning using smartphone camera," in *Proc. IEEE IPIN*, Sep. 2011, pp. 1–6.

[6] B. Xie, G. Tan, and T. He, "Spinlight: A high accuracy and robust light positioning system for indoor applications," in *Proc. ACM SenSys*, 2015, pp. 211–223.

[7] K. P. Subbu, B. Gozick, and R. Dantu, "Locateme: Magnetic-fields-based indoor localization using smartphones," *ACM Trans. Intell. Syst. Technol.*, vol. 4, no. 4, p. 73, 2013.

[8] K. Chintalapudi, A. Padmanabha Iyer, and V. N. Padmanabhan, "Indoor localization without the pain," in *Proc. ACM MOBICOM*, 2010, pp. 173–184.

[9] Q. Li, W. Li, W. Sun, J. Li, and Z. Liu, "Fingerprint and assistant nodes based Wi-Fi localization in complex indoor environment," *IEEE Access*, vol. 4, pp. 2993–3004, 2016.

[10] X. Wang, L. Gao, S. Mao, and S. Pandey, "DeepFi: Deep learning for indoor fingerprinting using channel state information," in *Proc. IEEE WCNC*, Mar. 2015, pp. 1666–1671.

[11] Y. Shu *et al.*, "Gradient-based fingerprinting for indoor localization and tracking," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2424–2433, Apr. 2016.

[12] H. Liu *et al.*, "Push the limit of WiFi based localization for smartphones," in *Proc. ACM MOBICOM*, 2012, pp. 305–316.

[13] M. Sugano, T. Kawazoe, Y. Ohta, and M. Murata, "Indoor localization system using RSSI measurement of wireless sensor network based on ZigBee standard," *Target*, vol. 538, p. 50, Jul. 2006.

[14] B. Yan and L. Xiaochun, "Research on UWB indoor positioning based on TDOA technique," in *Proc. IEEE ICEMI*, Aug. 2009, pp. 1-167–1-170.

[15] J. Wang, Y. Ma, Y. Zhao, and K. Liu, "A multipath mitigation localization algorithm based on MDS for passive UHF RFID," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1652–1655, Sep. 2015.

[16] S. S. Saab and Z. S. Nakad, "A standalone RFID indoor positioning system using passive tags," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1961–1970, May 2011.

[17] G. Lu, Y. Yan, A. Kolagunda, and C. Kambhamettu, "A fast 3D indoor-localization approach based on video queries," in *Proc. MMM*, 2016, pp. 218–230.

[18] H. Xu, Y. Ding, R. Wang, W. Shen, and P. Li, "A novel radio frequency identification three-dimensional indoor positioning system based on trilateral positioning algorithm," *J. Algorithms Comput. Technol.*, vol. 10, no. 3, pp. 158–168, 2016.

[19] W. Guan *et al.*, "A novel three-dimensional indoor positioning algorithm design based on visible light communication," *Opt. Commun.*, vol. 392, pp. 282–293, Jun. 2017.

[20] K. Wu, J. Xiao, Y. Yi, M. Gao, and L. M. Ni, "FILA: Fine-grained indoor localization," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 2210–2218.

[21] Z. Yang, C. Wu, and Y. Liu, "Locating in fingerprint space: Wireless indoor localization with little human intervention," in *Proc. ACM MOBICOM*, 2012, pp. 269–280.

[22] R. Faragher and R. Harle, "Location fingerprinting with Bluetooth low energy beacons," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2418–2428, Nov. 2015.

[23] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. ICML*, 2008, pp. 1096–1103.

[24] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[25] J. Wan *et al.*, "Deep learning for content-based image retrieval: A comprehensive study," in *Proc. 22nd ACM Int. Conf. Multimedia*, 2014, pp. 157–166.

[26] G. Hinton *et al.*, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.

[27] J. Li, Z. Struzik, L. Zhang, and A. Cichocki, "Feature learning from incomplete EEG with denoising autoencoder," *Neurocomputing*, vol. 165, pp. 23–31, Oct. 2015.

[28] T. Wang, D. J. Wu, A. Coates, and A. Y. Ng, "End-to-end text recognition with convolutional neural networks," in *Proc. IEEE ICPR*, Nov. 2012, pp. 3304–3308.

[29] C.-Y. Liou, W.-C. Cheng, J.-W. Liou, and D.-R. Liou, "Autoencoder for words," *Neurocomputing*, vol. 139, pp. 84–96, Sep. 2014.

[30] W. Zhang, K. Liu, W. Zhang, Y. Zhang, and J. Gu, "Deep neural networks for wireless localization in indoor and outdoor environments," *Neurocomputing*, vol. 194, pp. 279–287, Jun. 2016.

CHAO XIAO received the B.E. degree in remote sensing science and technology from Wuhan University, Hubei, China, in 2014, where he is currently pursuing the M.E. degree in photogrammetry and remote sensing. His research interests include indoor localization, intelligent video surveillance, and digital signal processing.

DAIQIN YANG (M'13) received the B.E. and M.E. degrees from the Huazhong University of Science and Technology, China, and the Ph.D. degree from The University of Hong Kong, all in electrical and electronic engineering. Prior to joining the Faculty at Wuhan University (WHU), China, in 2013, she spent several years with Philips Research Asia, Shanghai, Nanyang Technological University, Singapore, and the Institute for Info-comm Research, Singapore, as a Scientist and Research Fellow. She is currently an Associate Professor with the School of Remote Sensing and Information Engineering, WHU. Her research interests include sensor networks, intelligent systems, and remote sensing image processing.

ZHENZHONG CHEN (S'02–M'07–SM'15) received the B.Eng. degree from Huazhong University of Science and Technology and Ph.D. degree from Chinese University of Hong Kong, both in electrical engineering. Prior to joining Wuhan University (WHU), he was with MediaTek USA Inc., San Jose, CA, USA. He is currently a Professor and the Director of the Institute of Intelligent Sensing and Computing, WHU. His current research interests include visual perception, image and video processing, multimedia communications, and intelligent systems. He is a Board Member of VQEG, a Selection Committee Member of ITU Young Innovators Challenges, and a member of the IEEE Multimedia Systems and Applications Technical Committee. He was a recipient of the CUHK Young Scholar Dissertation Award, the CUHK Faculty of Engineering Outstanding Ph.D. Thesis Award, the Microsoft Fellowship, the ERCIM Alain Bensoussan Fellowship, and the First Class Prize of 2015 IEEE BigMM Challenge. He has served as an Area Chair or a TPC member of the IEEE ICME, ICC, and GLOBECOM. He has been the Special Session Chair of the IEEE World Forum of Internet of Things 2014 and the IEEE VCIP 2016 and the Publication Chair of the IEEE Conference on Multimedia and Expo 2014. He is a Co-Chair of the Immersive Media Working Group and a Co-Chair of the IEEE Multimedia Communication TC Networking Technologies for Multimedia Communication IG. He is an Editor of the *Journal of Visual Communication and Image Representation* and an Editor of the IEEE IoT Newsletter.

GUANG TAN (M'12) received the Ph.D. degree in computer science from the University of Warwick, Coventry, U.K., in 2007. From 2007 to 2010, he was a Post-Doctoral Researcher with INRIA-Rennes, Rennes, France. He is currently a professor with the Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, where he is involved in the area of distributed systems and networks. He is a member of the Association for Computing Machinery.

• • •