IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 10, 2018, accepted November 27, 2018. Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2018.2884193

Indoor Positioning Based on Fingerprint-Image and Deep Learning

WENHUA SHAO[1,2], (Student Member, IEEE), HAIYONG LUO[3,4], FANG ZHAO[1], YAN MA[2], (Member, IEEE), ZHONGLIANG ZHAO[5], AND ANTONINO CRIVELLO[6], (Student Member, IEEE)

[1]School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China
[2]Institute of Network Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China
[3]Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China
[4]Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China
[5]Institute of Computer Science, University of Bern, 3012 Bern, Switzerland
[6]Institute of Information Science and Technologies, Consiglio Nazionale delle Ricerche, 56124 Pisa, Italy

Corresponding authors: Haiyong Luo (yhluo@ict.ac.cn) and Zhongliang Zhao (zhao@inf.unibe.ch)

This work was supported in part by the National Key Research and Development Program under Grant 2018YFB0505200, in part by the National Natural Science Foundation of China under Grant 61872046, in part by the BUPT Excellent Ph.D. Students Foundation under Grant CX2017404, and in part by the Open Project of the Beijing Key Laboratory of Mobile Computing and Pervasive Device.

ABSTRACT Wi-Fi and magnetic field fingerprinting have been a hot topic in indoor positioning researches because of their ubiquity and location-related features. Wi-Fi signals can provide rough initial positions, and magnetic fields can further improve the positioning accuracies, therefore many researchers have tried to combine the two signals for high-accuracy indoor localization. Currently, state-of-the-art solutions design separate algorithms to process different indoor signals. Outputs of these algorithms are generally used as inputs of data fusion strategies. These methods rely on computationally expensive particle filters, labor-intensive feature analysis, and time-consuming parameter tuning to achieve better accuracies. Besides, particle filters need to estimate the moving directions of particles, limiting smartphone orientation to be stable, and aligned with the user's moving directions. In this paper, we adopted a convolutional neural network (CNN) to implement an accurate and orientation-free positioning system. Inspired by the state-of-the-art image classification methods, we design a novel hybrid location image using Wi-Fi and magnetic field fingerprints, and then a CNN is employed to classify the locations of the fingerprint images. In order to prevent the overfitting problem of the positioning CNN on limited training datasets, we also propose to divide the learning process into two steps to adopt proper learning strategies for different network branches. We show that the CNN solution is able to automatically learn location patterns, thus significantly lower the workforce burden of designing a localization system. Our experimental results convincingly reveal that the proposed positioning method achieves an accuracy of about 1 m under different smartphone orientations, users, and use patterns.

INDEX TERMS Indoor positioning, indoor localization, neural networks, fingerprint, feature extraction.

I. INTRODUCTION

In recent years, various location signals and related techniques have been studied to satisfy the increasing demand of indoor location-based service (LBS). Such signals include ultra-wideband (UWB), radio frequency identification (RFID), echo, Wi-Fi, and magnetic fields. UWB and RFID based schemes can achieve competitive accuracies, but they need to deploy base stations into the indoor environments [1], [2]. The accuracy of echo schemes are up to 23 cm, but the dense sampling points prevent it from continuous positioning in a large-scale deployment [3]. Nowadays, Wi-Fi networks are ubiquitous in indoor environments (e.g., shopping malls, offices, airports and railway stations). Therefore, Wi-Fi information can be used for personal indoor localization. However, Wi-Fi information can only provide rough position estimates due to its signal fluctuation, especially when moving pedestrians are considered [4]–[8]. Researchers also tried to use the Wi-Fi channel state information (CSI) [9], [10] to implement high precision positioning, but the method also needs to sample a large number of dense

training points, thus making it difficult for large-scale applications. Formed by the geomagnetic field and building's iron structures, the characteristics of the indoor magnetic field are ubiquitous and unique. Systems based on this signal can provide meter-level accuracies, but due to the low dimensionality of magnetic fields, it is difficult for them to provide global position estimates [11]–[14], [46], [47].

Considering the complementary features of Wi-Fi and magnetic field signals, researchers have tried to fuse together the two signals to implement high accuracy positioning systems that are available in large areas. For example, Shu *et al.* [12] fuse the RADAR [6] and magnetic field with particle filters, reaching 3m accuracy in almost the 80% of a large supermarket. In order to acquire the moving direction of particles, this method requires restricting the smartphone orientation aligned with the user moving direction, but this is not often easily detectable. Another drawback of the method is that the scheme needs to sample the training data for Wi-Fi and magnetic fields separately because of their different positioning principles. This aspect increases the cost of surveying the environment. Thus, in our work, we investigate whether deep learning—a recent and powerful machine learning paradigm—can provide a low surveying cost and device orientation free solution using available Wi-Fi and magnetic field signals. Considering the outstanding performance of deep learning in image classification, we construct a novel fingerprint image using Wi-Fi and magnetic signals to represent the features of locations. Then, these images are inputted into a deep CNN to predict real-time location estimates. The proposed method combines the Wi-Fi and magnetic field peculiarities in a single image; therefore, the two heterogeneous signals can be sampled at the same moment during a single process. On the other hand, the proposed fingerprint images are orientation-free, thus they remove the constraint due to the smartphone orientation alignment.

The main issue in applying CNN to fingerprint images is how to tackle the heterogeneity of the Wi-Fi and magnetic field signals. For example, compared with Wi-Fi signals, magnetic field signals are more stable, and with fewer dimensions and higher sampling rates. These differences make the feature channels of fingerprint images more complicated than ordinary images; therefore, it is difficult to learn the location patterns of these complex images with common convolution windows. In [15], we had partially described the effects of different convolution windows and neural network structures applied to the fingerprint images, and we proposed a Direct Acyclic Graph (DAG) neural network. The designed network adopts a two-level hierarchical structure. The neurons in the first level extract signal-independent patterns, then the second level organizes these extracted patterns into high-level images, analyze their patterns, and estimates final predictions.

Another challenge in creating positioning systems based on CNNs is to deal with the constraints of the network training. Traditional neural networks need a large training dataset to avoid overfitting problems. In order to achieve high positioning accuracies, the CNN based positioning system needs dense classification points. For example, if the target positioning accuracy is less than one meter, the distance of adjacent classification points should be less than that to provide the minimum resolution. However, gathering and sampling a large training dataset for such a great number of classification points is a formidable task because a real deployment cost would be too expensive. Therefore, in order to extract enough location features from a limited training dataset, we make full use of the high dimensionality of Wi-Fi signals and the stability of magnetic fields. We divide the training process into two steps. The first step leverages an early-stop method to prevent the overfitting of Wi-Fi features on small datasets. The second step executes a long-time training to learn sufficient features from the stable magnetic fields and the relations between the two signals.

As a conclusion, with respect to our previous work on this topic [15], the contributions of the paper are summarized as follows:

- We design a positioning system based on DAG CNN to overcome the stability difference of Wi-Fi and magnetic field signals. The proposed method leverages a hierarchical architecture to separately learn signal-independent and united feature patterns.
- Considering the limited training dataset of a large number of classification points, we extensively describe a two-step training method for the proposed positioning CNN to prevent the overfitting problem.
- According to [16], we test our system in a real test site, extending our preliminary previous results [15]. Experiments convincingly reveal that the proposed positioning method reaches remarkable performance also considering various smartphone orientations and different target users.

The rest of this paper is arranged as follows: section II introduces related works and highlights the differences between our proposal and the state-of-the-art systems, section III reviews the background and our empirical studies of magnetic field and Wi-Fi signals, section IV contains an overview of our system, section V, VI, VII and VIII extensively details the scheme of fingerprint construction, positioning model, network training and the fingerprint classification, section IX describes how the experimental campaign was set and conducted and, finally, section X describes conclusions and future works.

II. RELATED WORK

The request for efficient indoor LBS has spurred the development of many indoor positioning techniques, also due to growing commercial interests. The use of Wi-Fi and magnetic field signals has attracted continuous attention due to the ubiquity of wireless access points (APs) and magnetic field in indoor environments.

Traditional outdoor positioning systems require line-of-sight (LOS) measurements because the positioning schemes

rely on trilateration and triangulation. The performances of these approaches drop in indoor environments due to the frequent blockage of signals by walls and multipath propagation. Therefore, fingerprinting techniques are widely used in indoor LBS. After Bahl and Padmanabhan [6] presented the first Wi-Fi fingerprint positioning system RADAR, various Wi-Fi based algorithms have emerged to further improve the positioning accuracy and robustness. These algorithms can be divided into deterministic and probabilistic methods. Deterministic methods leverage various similarity metrics to differentiate online and offline fingerprints, including Euclidean distance [17], Tanimato similarity [18], and cosine similarity [19], or using an ensemble of different distances [45]. The location with the most similar fingerprint in signal space is considered as the positioning result. Compared with probabilistic methods, deterministic methods are easier to implement, but the probabilistic methods show better localization accuracy. Probabilistic methods use statistical inference to measure the difference between real-time signal fingerprints and fingerprint models. These algorithms find the target's location by evaluating the maximum likelihoods from training datasets. For example, Horus [20] estimates user locations using probabilistic models that are able to reflect signal distributions in the site. Centaur's [21] Bayesian network, Ouyang *et al.*'s [22] expectation-maximization and Xiao *et al.*'s [23] conditional random field method achieve decent accuracies through probabilistic inference. Probabilistic algorithms rely on distribution assumptions, such as Gaussian noise or independence, therefore, the training of probabilistic models is more complicated and it requires more datasets than deterministic algorithms. Because of the random fluctuation of Wi-Fi signals, the performances of Wi-Fi based algorithms are difficult to achieve high positioning accuracies, especially for moving targets.

It is reported that Wi-Fi based schemes achieve about 5-10 meter positioning accuracies; on the contrary, geomagnetism schemes can achieve a much higher accuracy (around 1-5 meters) [24]. There are several approaches to harness the indoor magnetic field for positioning purpose. Wang *et al.* [25] extract magnetic anomalies with unsupervised learning, then exploit these anomalies to pinpoint specific indoor locations as landmarks. Although these kinds of methods are simple and efficient to implement, their magnetic field information are not fully utilized. Magicol [12] models temporal magnetic measurements as sequences of strings, then solve target positions with traditional string matching and dynamic programming methods. However, these dynamic programming based matching schemes are computationally expensive. Shu *et al.* [26] propose a leader-follower mode based scheme. Such a mode is particularly useful when floor plans are missing, but this mode is tailored to some specific applications. Magnetic field based schemes are found to achieve reasonably high localization accuracy with low energy consumption, but it is difficult for them to provide large-scale site deployments due to the decreasing discernibility of magnetic fingerprints [27]. Therefore additional sensing information is needed to make the fingerprints more discernable [13].

Besides deterministic and probabilistic methods, nowadays, the literature offers different approaches. Deep learning is a fast-growing field in machine learning due to the development of massive amounts of computational power. Researchers have used deep learning in different contexts and applications: surveillance camera to localize vehicles, and channel state information (CSI) to localize laptops [9], [28], [29]. Vieira *et al.* [30] investigate the application of CNNs for fingerprint-based positioning using massive MIMO channels on the COST 2100 model. A group fuses cameras and magnetic fields with the neural network, but this method restricts phone's attitude to be upright to use the camera [31]. Nowicki and Wietrzykowski [32] investigate using the deep neural network to classify buildings and floors. Researchers also try to implement indoor positioning with backpropagation neural networks (BPNNs) based on Wi-Fi fingerprint, however, this method needs massive fingerprint sampling work [33]. As a summary, current deep learning based schemes focus on predicting positioning results with single signal resource, which limits their accuracy and scalability.

Different from aforementioned deep learning based positioning methods, our proposed algorithm requires no dedicated hardware (camera or antenna arrays) or specific CSI signal. We introduce deep learning method to localize the commodity smartphones. Several ubiquitous signals including Wi-Fi and magnetic field signals are used, which is free of privacy issues. Therefore, our approach is more pervasive and cost-effective.

For the implementations, engineers have developed various platforms for deep learning based researches, including MatConvNet [34], Tensorflow [35], and Caffe [36]. The system proposed in this paper is built on top of MatConvNet.

As a summary, compared to current positioning systems, the proposed method combines Wi-Fi and magnetic signals in a common fingerprint image, therefore the two location fingerprints can be sampled at the same time, reducing sampling cost. On the other hand, our system requires no orientation information, and consequently, it has no restrictions on smartphone attitudes. Therefore, users are able to use their smartphones as usual. To these extents, our system is more practical than existing Wi-Fi and magnetic field based positioning systems.

III. FEATURES OF INDOOR MAGNETIC FIELD AND Wi-Fi SIGNALS

Indoor magnetic field signals are relatively stable for a long-term and have remarkable spatial discernibility in the local area, but their discernibility decrease when a larger area is considered [27], [37], [38]. For instance, Fig. 1. illustrates two magnetic fingerprints gathered along the same corridor in March and in May, respectively. The similarity of the two fingerprints reflects the stability of the indoor magnetic field signals. For the local discernibility example, let us assume

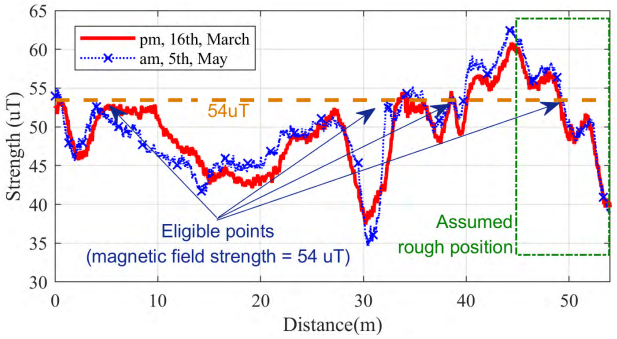


FIGURE 1. Stability of indoor magnetic field strength. The two fingerprint traces have been gathered along a 52-meter corridor on two different days, two months apart.

that a user measures a magnetic strength of $54\mu T$. If we have a priori knowledge that the user's rough position has a distance of $45 \sim 55$ meters from the zero point, we can refine positioning results by comparing magnetic strengths, and estimate that the user is located approximately 50 meters from the zero point with an error less than one meter. However, without this local area prerequisite, we have to average all eligible positions with the magnetic strength of $54\mu T$. This averaging method decreases positioning accuracy.

The uniqueness of Wi-Fi media access control (MAC) address and its limited signal coverage range enable Wi-Fi signals being used in positioning systems. However, Wi-Fi based system is not able to reach high-accuracy performances when the target is represented by a moving smartphone. In fact, Wi-Fi signals are susceptible to surrounding electromagnetic noises like transmission rate and power adaptions [39], causing signal strength fluctuates over the time. Thus, Wi-Fi based system generally confuse nearby positions. Furthermore, the signal scanning process takes several seconds to traverse all wireless channels; therefore, users may have walked several meters from the sampling point to the end of the scanning process. This factor is critical especially for statistics-based positioning methods [5], [40] that need several scanning periods to collect enough samples on the same place to evaluate signal distributions.

IV. SYSTEM ARCHITECTURE

As Fig. 2 shows, the proposed positioning system is comprised of five functional modules: the CNN positioning, CNN training, fingerprint-image classification, fingerprint-image construction, and step detector module. In addition, the position label is a key-value pair (KVP) containing a position ID and its coordinates. The position label database is a set of KVP uniformly scattered in the testing area. The neural network (NN) parameter database stores the CNN positioning model for different positioning sites. The sensor data used in our system include Wi-Fi, magnetic field, and acceleration.

A. FINGERPRINT CLASSIFICATION

In order to construct the mapping between environmental signals and locations. Users collect location signals

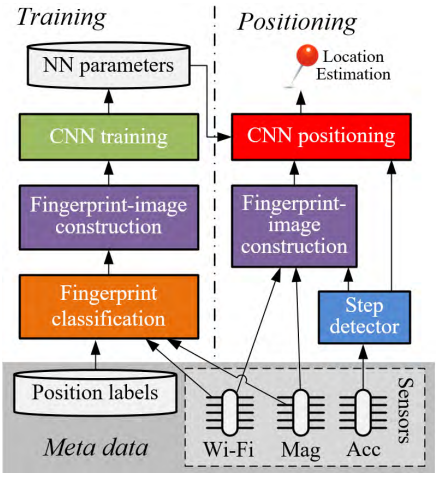


FIGURE 2. The overall architecture of our proposed indoor positioning system based on convolutional neural network and fingerprint image learning. Acc means a triaxial accelerometer. Mag means a triaxial magnetometer. Wi-Fi represents the Wi-Fi chip. CNN means a convolutional neural network.

(including Wi-Fi and magnetic field data) walking along sampling lines. Then the fingerprint classification module cut fingerprint sequence in similar length for every position labels. The rule is that the distance between fingerprint segments and position labels should be small enough.

B. FINGERPRINT-IMAGE CONSTRUCTION

Considering the difference in data length and the number of AP in each fingerprint segments, the fingerprint-image construction module normalize all the fingerprints in the same data length and AP set. This module is a common module both in the off-line training and in the online positioning procedures. The only difference is that the fingerprint-image in the training phase is labeled and the positioning phase is not.

C. CNN POSITIONING

We use a CNN algorithm to compute the user positions. Because we have discretized indoor locations into position labels, therefore the result is the ID of these labels. The positioning CNN is customized to support the different stability and the number feature channels of Wi-Fi and magnetic field data in the fingerprint image.

D. CNN TRAINING

The positioning CNN is able to automatically learn location features. However, the training dataset for each position labels is rather small, leading the algorithm to overfitting problems. Therefore, the CNN training module divides the learning process into two steps in order to adopt different learning strategies on the Wi-Fi branch and the rest of the neural network. Then, the trained NN parameter is stored in a database.

E. STEP DETECTOR

The positioning result is updated at each step. Every time a new step is reported by the step detector, a new

fingerprint-image is generated by the fingerprint-image construction module, triggering the CNN positioning module to estimate user locations. The step detector is implemented by forming stride templates offline and applying deep learning classification with the accelerometer signal online to detect real-time step events [41].

V. FINGERPRINT-IMAGE CONSTRUCTION

Different from the traditional feature extraction manners of position signals, we propose a novel image-based method. Sensor data is a time series containing multiple feature channels. In fact, an ordinary image is a three-dimensional matrix, defined by width, height and color channels (red, green, and blue). Therefore, in order to convert raw sensor data into the location images, we reconstruct the sensor data series. We use the first dimension of fingerprint images to represent the length of the measurements gathered from the sensors, and the second dimension to store multi-feature channels, in other words, image width and height respectively. The third dimension is equal to one, therefore it is omitted.

In order to use images as inputs of the proposed CNN, the image dimensions have to be normalized to the same size. The fingerprint image F is composed of magnetic-fingerprint part F^m and Wi-Fi fingerprint part F^w, represented as:

$$F^m = [MS_1, MS_2, \cdots, MS_n] \tag{1}$$

$$F^w = \begin{bmatrix} RSSI_{11} & RSSI_{12} & \cdots & RSSI_{1n} \\ RSSI_{21} & RSSI_{22} & \cdots & RSSI_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ RSSI_{m1} & RSSI_{m2} & \cdots & RSSI_{mn} \end{bmatrix} \tag{2}$$

$$F = \begin{bmatrix} (F^m - \mu(F^m))/\sigma(F^m) \\ (F^w - \mu(F^w))/\sigma(F^w) \end{bmatrix} \tag{3}$$

where F^m and F^w are resampled to fit the image with the normalized image width n and MS means magnetic strength. In order to accelerate the convergence of the training process, the fingerprint image F normalizes the F^m and F^w by deducting the mean and dividing the standard deviation respectively. n is defined by the detection range r and the storage density ρ as shown in formula (4). The balanced structure guarantees the classification point lies in the center of the fingerprint image.

$$n = 2 \cdot round(r \cdot \rho) + 1 \tag{4}$$

The height normalization process of the fingerprint images is different between the magnetic field and Wi-Fi signals. The magnetic field part only uses strengths as location features, so the magnetic strengths F^m are stored as a $1 \times n$ vector. The Wi-Fi fingerprint image is stored as a $m \times n$ matrix, where the height m is equal to the number of all Wi-Fi MAC addresses detected in the positioning area. Elements in F^w are equals to the received signals strength (RSS) collected from APs. In order to keep the consistency of fingerprint images, the RSS of an AP at point P_x whose strength is not measured is set equal to -120dB. Finally, the final image is built by concatenating magnetic field and Wi-Fi parts along the second dimension.

VI. CNN POSITIONING MODEL

The proposed positioning solution combines the location features of Wi-Fi and magnetic field fingerprints using a hierarchical CNN. The network is built using the MatConvNet library that has provided ready-to-use modules. In this section, we introduce the algorithm architecture and we describe modules used in our system. Furthermore, we analyze the network attributes obtained from Wi-Fi and magnetic field signals, then we present the final design of the positioning CNN.

A. BASIC LAYERS OF THE PROPOSED POSITIONING CNN

The proposed positioning CNN has a stacked layer structure, as Fig. 3 reveals. The convolution, pooling, rectified linear unit (ReLu) and the fully connected multilayer perceptron (MLP) modules play an important role in learning location features from multiple fingerprint images.

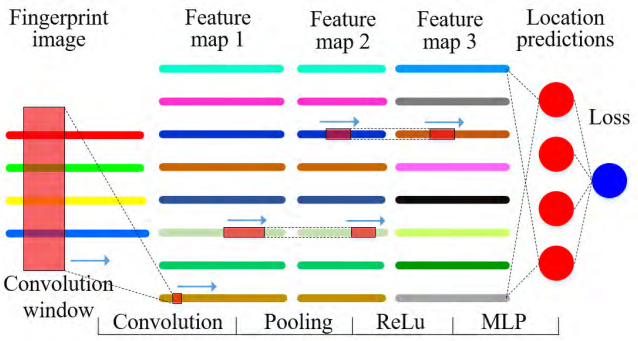


FIGURE 3. A basic CNN structure of the proposed positioning system.

The convolution layer has H_1 random initialized convolution windows. In order to learn location features, these convolution windows slide along the input fingerprint image with given moving stride S_1. For each stride, the convolution window calculates the dot product of the convolution window vector and the image part vector under the convolution window. The length of the convolution window is W_1. The outputs of the convolution layer is H_1 new feature maps.

$$\begin{cases} y_{i_2,k_2} = b_{k_2} + \sum_{i_1=1}^{W_1} \sum_{j_1=1}^{m} \omega_{i_1,j_1} \cdot F^f_{i_2 \cdot S_1 + i_1, j_1, k_2} \\ 1 \le i_2 \le 1 + \left\lfloor \dfrac{n - W_1}{S^1} \right\rfloor \\ 1 \le k_2 \le H_1 \\ F^f \in \{F^m, F^w\} \end{cases} \tag{5}$$

where ω_{i_1,j_1} is an element of the convolution windows. $F^f_{i_2+i_1,j_1,k_2}$ is an element of the fingerprint images, and b_{k_2} is the bias of a convolution window.

The pooling layer down-samples the feature maps generated by the previous layer to reduce the data size and keep

the main location features. In the proposed system, the max-pooling layer is applied. It leverages a max filter to the sub-regions of the initial feature map and takes the most obvious feature of that region, creating a new output feature maps. Each element of this new output is the most obvious feature of a region considering the original input. This process is represented as:

$$\begin{cases} y_{i_4,k_4} = \max\limits_{1\le i_3\le W_3} y_{i_4\cdot S_3+i_3,k_4} \\ 1 \le i_4 \le 1 + \left\lfloor \dfrac{\max(i_2) - W_3}{S^3} \right\rfloor \\ 1 \le k_4 \le \max(k_2) \end{cases} \tag{6}$$

where W_3 and S_3 are the windows width and the windows stride of the max pooling filter, respectively.

The ReLu layer is a nonlinear activation function, which adds potential non-linearity to the mapping between features and locations, represented as:

$$y_{i_5,k_5} = \max\left(0, y_{i_4,k_4}\right) \tag{7}$$

Finally, the fully connected MLP calculates prediction values of all position labels. This layer is a special case of the convolutional layer when the convolutional window width $W_6 = \max(i_5)$. The number of convolutional windows H_6 equals to the number of classification positions C. The output of this layer is a prediction vector $y_{i_6}(i_6 = 1 : H_6)$. Every element in this vector is a score of the predicted positions. The estimated location is the position with the highest score.

In the training phase, the CNN scheme tunes the neural network by comparing the prediction and the ground-truth position c. In order to evaluate the prediction errors, the proposed system adopts the soft-max loss function. The soft-max loss is calculated as:

$$\ell = -\log \frac{e^{y_c}}{\sum_{i_6=1}^{C} e^{y_{i_6}}} \tag{8}$$

where y_c is the prediction score of the ground truth label.

B. Wi-Fi BRANCH

In order to reduce the influence of the signal heterogeneity, the proposed positioning CNN leverages a Wi-Fi and magnetic field branch to separately learn high-level location features that are signal independent. Then, a united branch is used to learn the relation between ground-truth positions and high-level features to predict location estimates.

The biggest challenge of developing Wi-Fi based positioning is the random fluctuation of Wi-Fi signals. Therefore, researchers increase the number of sample time to find the statistical features of APs, but this solution increases sampling workloads and does not work well for moving targets. In order to localize pedestrians with less AP samples, we use long convolution windows and small size neural networks to prevent the potential overfitting problems.

Convolution windows are able to automatically extract location features from Wi-Fi fingerprint images. As Fig. 4 shows, Wi-Fi fingerprint images, and convolution windows

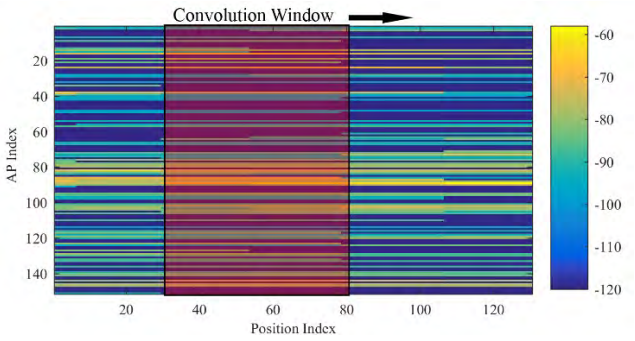


FIGURE 4. Wi-Fi fingerprint image and convolution window. The horizontal stripes are RSSs of APs.

can be represented as two matrixes. The convolution matrix is randomly initialized and slides along the fingerprint image matrix to extract location features. For each training epoch, the convolution layer generates a new feature map to represent high-level location features, and then the network leverages backpropagation algorithms to update the parameters of the convolution matrix. Fig. 5 illustrates a convolution window before and after the neural network training. Although this filter is randomly initialized, it produces a strong and clear pattern after the learning process, as depicted in Fig. 5(b). Therefore, with the learned patterns in the training phase, the neural network is able to classify real-time fingerprint images during the positioning phase.

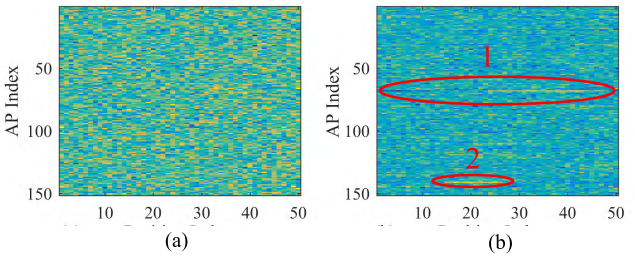


FIGURE 5. Weights of a convolution window for Wi-Fi positioning branch. Figure 5(a) shows the initial weight state of a convolution window. The windows are randomly initialized. Figure 5(b) shows the convolution window after the training phase. The two lines in circle 1 and circle 2 show the features learned from AP 68 and AP 141.

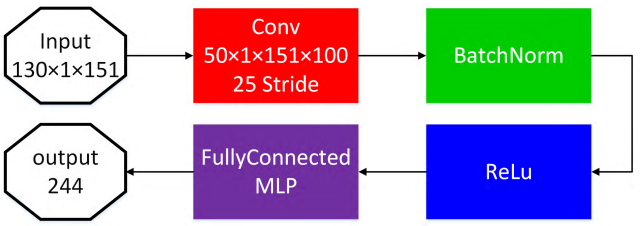


FIGURE 6. The proposed Wi-Fi positioning branch applying to an example of our system.

Fig. 6 specifies the structure of the long convolution windows and the small-size neural network. The proposed CNN uses one convolution layer to extract location features from

the inputted fingerprint images, followed by a batch norm layer to accelerate convergence, and then a ReLu layer to add nonlinearity, finally the MLP converts feature maps to location labels. Based on our experiments, we set the length of the fingerprint images as 10.4 meters, approximately equal to the length of 16 steps (about 0.65 meters for one-step length). The length of convolution windows is set to 4 meters, with 50% overlap between different windows. Considering the storage density of Wi-Fi samples equal to 12.5 samples/meter, the input 10.4 meter fingerprint image corresponds to 130 samples, and the 4 meters convolution window comprises of 50 samples as shown in Fig. 6. The height of the fingerprint images and the convolution windows are equal to the number of access points within the positioning area (151 AP addresses in our experiments). The output of the neural network is a weight vector of all predictions, consequently is equal to the number of classification points (244 points in our experiments). In conclusion, the proposed network decreases the network freedom, therefore relieves the overfitting problem.

C. MAGNETIC FIELD BRANCH

Magnetic field features are more stable and complex than the previously described Wi-Fi features. In order to learn these feature patterns, we propose a deeper CNN structure with smaller convolution windows. Both the small window size and deep CNN structure can increase network freedoms, therefore improve the learning ability for complex position patterns. As Fig. 7 shows, the proposed network consists of nine layers: three convolution layers for learning features, three batch normalization layers for accelerating the training convergence, one max-pooling layer for reducing data size, one ReLu layer for adding nonlinearity, and one fully connected MLP layer for outputting final predictions.

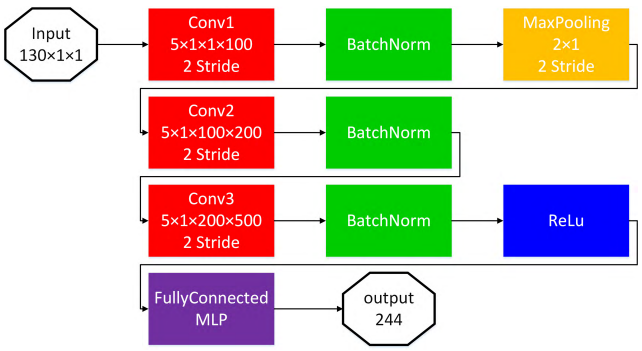


FIGURE 7. The proposed magnetic positioning branch applying to an example of our experiments.

The proposed magnetic CNN learns feature pattern by adjusting the bias of convolution windows. The convolution window consists of two kinds of parameters: weights and bias. During the training phase of the Wi-Fi positioning branch, the network mainly tunes the weight. However, considering the magnetic field branch, the network mainly tunes

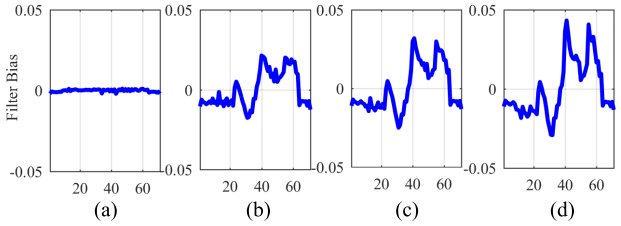


FIGURE 8. Bias values of a convolution window for magnetic field based positioning branch. We run the training program for 60 epochs, the figure shows four epochs as an example. The four bias vectors are evaluated during the last convolution layer. We trim the first 70 biases from the whole 244 biases for clarity.

the bias. As Fig. 8 shows, the bias patterns become gradually stronger as the training proceeds. The network is learning the location pattern from magnetic field fingerprints. Magnetic strength is stable, so its bias has a constant pattern. Compared with Wi-Fi branch, the convolution window size of the magnetic field is smaller and the number is larger. Therefore, hundreds of random initialized magnetic convolution window weights have covered almost all possible patterns. Hence, with respect to the magnetic fingerprint images, the main differences between them are the combination of convolution window biases, rather than window weights.

D. UNITED BRANCH

In order to take the advantages both of the magnetic field and Wi-Fi fingerprint images, this paper compares two kinds of united positioning methods: the direct and the branch method. The direct method uses convolution windows sliding across the united fingerprint image and produces predictions as outputs. The branch method calculates separate positioning results with the proposed Wi-Fi and magnetic field branches, and then these results are used as inputs of the united branch. Finally, the united branch is responsible for outputting predictions in terms of positioning evaluations. As Fig. 9 shows, the united branch method achieves the best positioning results compared with the performances achieved by the Wi-Fi branch only or the magnetic field branch only. Furthermore, the performance of the direct method is the worst, because the common convolution windows are not able to simultaneously capture features from both Wi-Fi and magnetic field fingerprint parts.

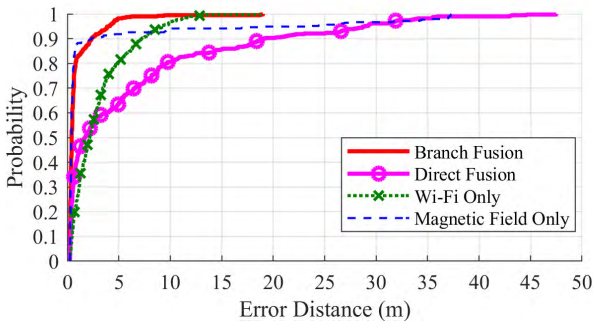


FIGURE 9. Performances comparison using different localization methods.

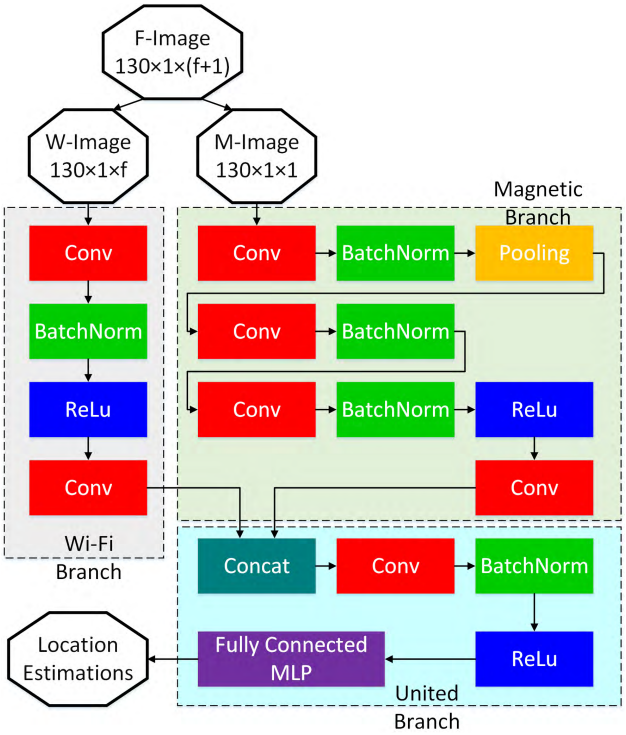


FIGURE 10. The architecture of our proposed CNN. Our system consists of three branches: the Wi-Fi branch, the magnetic branch, and the union branch.

As Fig. 9 shows, the maximum Wi-Fi branch error is equal to 18 meters. Considering indoor environments, these results are not high. On the contrary, the magnetic field branch can easily achieve high positioning accuracy, but it also suffers from big jitter error problems. The proposed positioning CNN takes the advantages from both Wi-Fi and magnetic branches. As Fig. 10 reveals that the proposed CNN leverages a DAG structure to combine the Wi-Fi and magnetic branches. United fingerprint images are decomposed into separated Wi-Fi and magnetic field parts. Then, the Wi-Fi branch produces a high-level signal-independent prediction vector as intermediate outputs. Namely, it is a vector whose elements are the prediction scores of the classification positions. Similarly, the magnetic field branch produces a second prediction vector. Successively, the united branch concatenates the Wi-Fi and magnetic field prediction vectors, then uses the combined prediction vector as the input of the united branch to find the patterns between ground-truth positions and the combined prediction vector, and finally estimates the user locations.

E. TIME COMPLEXITY ANALYSIS

The key parameters affecting the time complexity of the proposed positioning CNN is the number of APs M, the length of fingerprint features N, the number of convolution windows in Wi-Fi branch H_w, the number of convolution windows in magnetic field branch H_m, the number of convolution windows in united branch H_u, and the number of classification points P.

Wi-Fi branch time complexity: considering the convolution layer is the multiplication of Wi-Fi fingerprint images and filter matrices, based on formula (5), the time complexity of the convolution layer is $O(H_w \cdot M \cdot N)$. The time complexity of the batch normalization and ReLu layer is $O(M \cdot N)$. The time complexity of the MLP layer is $O(P \cdot M \cdot N)$. Therefore, the total complexity of the Wi-Fi branch is $O(H_w \cdot M \cdot N) + O(M \cdot N) + O(P \cdot M \cdot N) \sim O((H_w + P) \cdot M \cdot N)$.

Magnetic field branch time complexity: the convolution process of the magnetic field branch is the multiplication of magnetic field fingerprint image and filter matrices, the time complexity of convolution layers is $O(H_m \cdot N)$. The time complexity of the pooling, batch normalization and ReLu layer is $O(N)$. The time complexity of the MLP layer is $O(P \cdot N)$. Therefore, the total complexity of the magnetic field branch is $O(H_m \cdot N) + O(N) + O(P \cdot N) \sim O((H_m + P) \cdot N)$.

United branch time complexity: the convolution process of the united branch is the multiplication of concatenated images and filter matrices, the time complexity of convolution layers is $O(H_u \cdot N)$. The time complexity of the concatenation, batch normalization and ReLu layer is $O(N)$. The time complexity of the MLP layer is $O(P \cdot N)$. Therefore, the total complexity of the united branch is $O(H_u \cdot N) + O(N) + O(P \cdot N) \sim O((H_u + P) \cdot N)$.

Positioning CNN time complexity: in conclusion, the overall time complexity of the positioning CNN is $O((H_w + P) \cdot M \cdot N) + O((H_m + P) \cdot N) + O((H_u + P) \cdot N) \sim O((H_w \cdot M + P \cdot H_m + 2 \cdot P + H_m + H_u) \cdot N)$. The order of magnitudes of H_w, H_m, H_u, M, N and P are the same, therefore the overall time complexity can be simplified as $O(N^3)$. Considering the length of fingerprint features N is limited, the calculation complexity of our proposed algorithm is acceptable.

VII. CNN TRAINING

Considering that training methods influence the performance of the network, we present the training method for every single branch, and we propose the early-stop assisting training method.

The number of training epochs influences the performance of neural networks. Considering that the Wi-Fi and magnetic branches have different features, we use different training methods to tune the parameters of different branches.

Wi-Fi branch training: Fig. 11 illustrates that the classification error rate of the Wi-Fi branch in different learning epochs. If the training error rate drops continuously after each training epoch, it indicates that the network keeps learning features from training data. The top-one validation error rate has little change in Fig. 11(a), because the fluctuation of Wi-Fi signals makes confused classification between adjacent points. Consequently, the top-one error rate is not able to reflect the learning process of the Wi-Fi branch when the validation dataset is analyzed. Fig. 11(b) shows the trend of the error rate, considering the top-15 error rate.

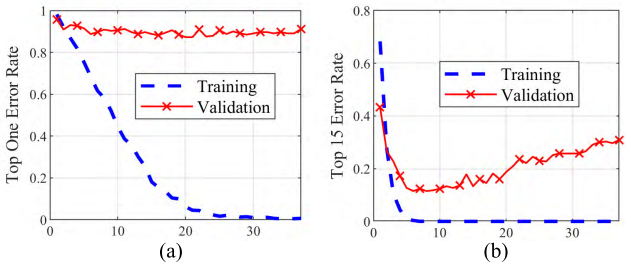


FIGURE 11. The classification error rate of Wi-Fi branch training. The blue lines show the error rate during the training phase. The red lines are related to the validation phase.

During the validation phase, the error rate drops at the beginning. When over-training occurs (after the sixth training epoch), the error rate begins to rise. Therefore, we propose the early-stop method that fully trains the Wi-Fi positioning branch until minimum top-15 error rate is determined, then chooses the CNN model at the minimum epoch as the final model.

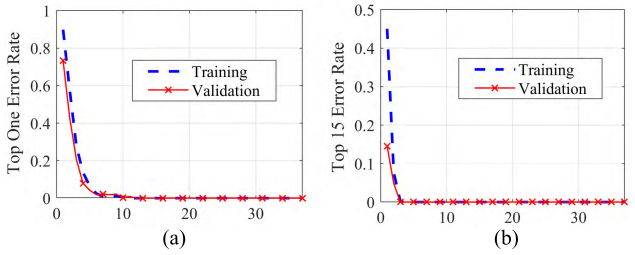


FIGURE 12. The classification error rate of the magnetic field branch during the training phase. The blue lines represent the error rate during the training phase. The red lines are related to the validation phase.

Magnetic field branch training: indoor magnetic fields are more stable than Wi-Fi signals, so magnetic field branch performance improves as the training process continues. As Fig. 12 shows, when the training phase starts, both the training and the validation error rate drops continuously. The simultaneous drop indicates the high similarity between training data and validation data. Therefore, we suggest continuing the magnetic field training phase until the error-rate keeps stable.

The early-stop assisting training method for the united positioning CNN: considering the multi-branch structure of the proposed positioning CNN, this paper divides the training procedure into two steps, as Algorithm 3 shows. The Wi-Fi branch training phase is considered separately. The magnetic field and the union branch training phase runs simultaneously.

VIII. FINGERPRINT CLASSIFICATION

The proposed system leverages classification methods to localize real-time positioning fingerprints. Therefore, the system discretizes continuous indoor positions into separate classification points. In order to improve positioning accuracy, the classification points should be dense enough. For example, in order to achieve meter-level positioning, our experiment set the density higher than one classification point per square meter.

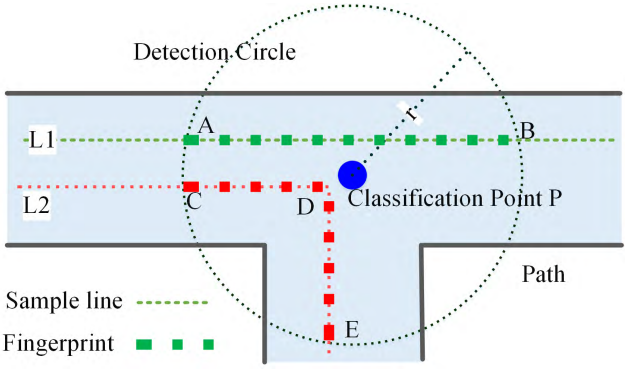


FIGURE 13. Fingerprints within the detection circle are labeled belonging to point P. As light gray is shown a T shaped corridor. r is the radius of the detection range.

As the Algorithm 2 reveals, features from nearby classification points are sampled in order to be utilized as inputs in the neural network-training phase. In our practice, a tester walks at a uniform speed (∼1 m/s), following a sample line, and holding a smartphone. As Fig. 13 shows, the sample line could be a straight line (L1), or a polygonal line (L2) along with a corridor intersection. Then, these raw samples are classified into different classification points. All fingerprint-images are required to have similar lengths. Given a fingerprint length l, we extracts feature segments with the help of a detection circle. The radius r of the detection circle is half of the fingerprint length l. Then feature segments within the circle are cut out as candidate segments. All the candidate segments shorter than the 80% (based on our experiments) of the diameter are discarded because they are considered too far from the current classification point to convincingly represent its features. Then, the rest segments are labeled as fingerprints of the classification point P.

IX. EXPERIMENTS AND RESULTS

We evaluate the key coefficients used in our proposed CNN based localization algorithm, and we perform further applicable tests considering different devices, users, and use cases. Finally, we compare our presented solution with other state-of-the-art positioning methods.

A. SETUP OF THE EXPERIMENTAL CAMPAIGN

The system is tested on the seventh floor of the Research Building of the Institute of Computing Technology, Chinese Academy of Sciences. This representative office building covers an area of $60m \times 40m$, and have fourteen floors, each with ceiling heights of 3 meters. The test site is comprised of office rooms and workstations, as Fig. 14 reveals. Workstations are made of wood and plastics, and the heights of workstations are $1.5m$. There are few iron filing cabinets along the paths. The test site includes two scenarios: corridors and working areas. Different from the working areas, corridors

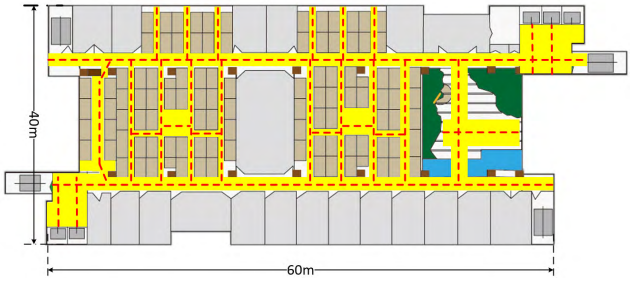


FIGURE 14. The floor plan of the test site. The yellow zone represents reachable areas. The red dash lines stand for data sampling routes. The gray blocks represent office rooms. The orange blocks represent workstations.

are near walls and pillars, therefore the magnetic field features gathered are distinctly different. All the reachable areas are discretized into 533 classification points. In the sampling phase, users collect data four times along all the sample lines. In the positioning phase, testers walk inside the reachable area and make marks when reaching the testing points. Finally, the error distances between the positioning and the testing points are calculated with 2D Euclidean distance. The testing paths are longer than 200 meters.

The positioning system is implemented on a cloud-computing platform, because the cloud platform has high flexibility in testing new versions of positioning algorithms and more powerful data-processing capability, compared to smartphone platforms. It takes about 2 milliseconds to calculate positioning results for each positioning request.

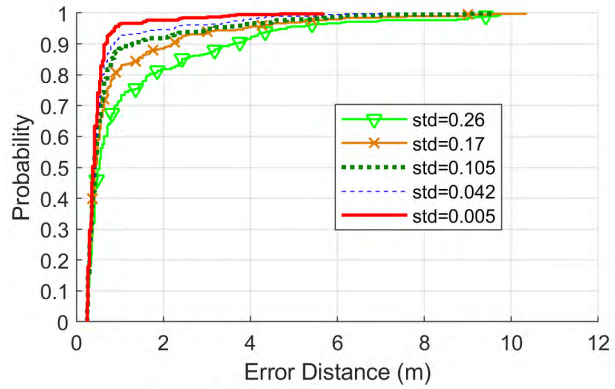


FIGURE 15. Positioning performance of different early-stop parameters.

B. THRESHOLD INFLUENCE OF THE TRAINING-STOP

In order to evaluate the training-stop parameter ε of the united positioning CNN, we compare the accuracy using different training-stop thresholds. In Algorithm 1, after training the Wi-Fi branch and keeping it unchanged, we train the magnetic and the united branch. For each training epoch, a sliding window with five adjacent top-one error rates is leveraged to calculate the standard deviation. As Fig. 15 reveals, the positioning performance increases as the standard deviations of learning rates drop, which introduces more training cost.

Algorithm 1 Early-Stop Training of the United Positioning CNN

Initialization:
Initializing united positioning CNN with random coefficients.
Wi-Fi branch training:
Isolating Wi-Fi branch.
Inputting training data and labels.
for each training epoch **do**
  Calculating the top-15 error rate and epoch index.
  Updating the minimum top-15 error rate and epoch index.
  Updating the training model with the minimum top-15 error rate.
  if current epoch – minimum top-15 epoch index > N **do**
    Breaking for loop
  end if
end for
Replacing the Wi-Fi branch of the united CNN with the trained Wi-Fi model
Setting the learning rate of the Wi-Fi branch to be zero.
Magnetic and united branches training
Inputting training data and labels.
for each training epoch **do**
  Calculating the standard deviation e of the last few top one error rates
  if $e < \varepsilon$ **do**
    Returning the current CNN model.
  end if
end for

When standard deviation is less than 0.042, the performance improvement brought from the training becomes less significant. To balance the training cost and performance, we choose 0.042 as the typical value of the training-stop parameter ε in this paper.

C. INFLUENCE OF DIFFERENT FINGERPRINT LENGTHS

In this experiment, we evaluate the Influence of different fingerprint lengths on positioning accuracies. We separately test positioning accuracies with different lengths using Wi-Fi branch only, magnetic field branch only and the united CNN, respectively. Our system is triggered by step events, and we measure fingerprint lengths with strides (2 steps, ∼1.3m). In this experiment, performed along a single corridor, we compare the positioning performances with fingerprint lengths ranging from 1 stride to 20 strides.

1) Wi-Fi BRANCH

Fig. 16 shows that longer Wi-Fi fingerprints generally achieve better positioning accuracies. When fingerprint images are

Algorithm 2 Procedure of the Fingerprint Classification

1: Uniformly generating N_p position labels into the testing area
2: Considering as input the begin and the endpoints of sampling lines
3: Sampling Wi-Fi and magnetic data along the sampling lines
4: **for** each position label $p_i = (x_i, y_i)\ (1 \leq i \leq N_p)$ **do**
5: **for** each sampling lines **do**
6: **if** $\exists$ *a line segment* l near p_i **do**
7: Constructing a fingerprint image F with the sampled data in l
8: Labeling F with p_i
9: Adding F to training set S
10: **end if**
11: **end for**
12: **end for**

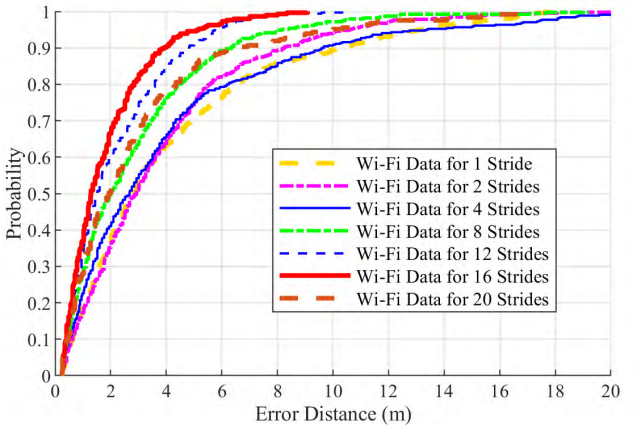


FIGURE 16. Positioning performances of Wi-Fi positioning branch using different strides.

too short (e.g., from 1 to 4 strides) the obtained positioning accuracies are low. A Wi-Fi scan takes about $2 \sim 3$ seconds to finish, which roughly corresponds to the length of four steps. Therefore, the fingerprint images with less than four steps have the same features and, consequently, the branch shows the same performance. On the other hand, when fingerprint images are too long, e.g. 20 strides, the positioning accuracy drops to the same level of eight strides. In fact, the 20 stride fingerprint images cover a path longer than 13 meters, and it contains too many features of adjacent locations. In conclusion, the length of the Wi-Fi fingerprint image should be neither too short nor too long.

2) MAGNETIC FIELD BRANCH

Fig. 17 shows, the performance of the magnetic branch increases as the length of magnetic images increases. Magnetic field signals are more stable than Wi-Fi signals, which is helpful in precisely align the real-time and the model

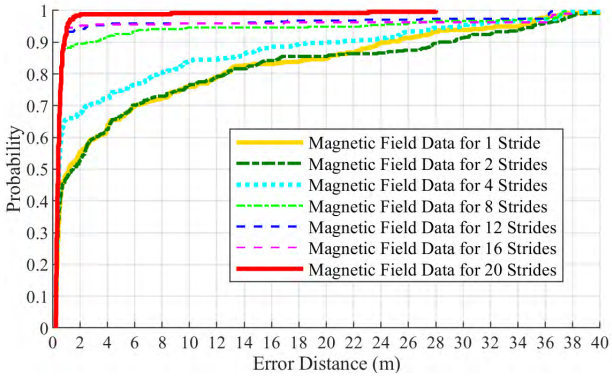


FIGURE 17. Positioning performances of magnetic positioning branch with different strides.

fingerprint images. Comparing Fig. 17 with Fig. 16, we can find that the positioning errors produced by the Wi-Fi branch are concentrated around the ground-truth location and the errors of the magnetic branch are distributed around the whole positioning area.

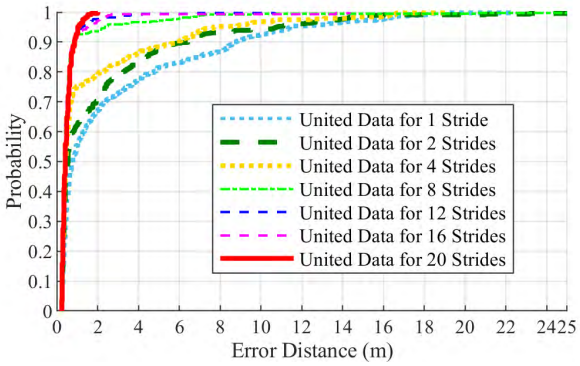


FIGURE 18. Positioning performances of the united branch of the positioning CNN with different strides.

3) UNITED BRANCH

Fig. 18 illustrates that the united branch of the positioning CNN performs better than the single Wi-Fi and magnetic field branches. With respect of the lowest errors, the united branch performs as good as magnetic field branch. On the other hand, the largest errors are limited to the same level as the Wi-Fi branch. The experimental result convincingly reveals that the proposed positioning CNN successfully gathers the advantages of the Wi-Fi and magnetic field signals. It is worth noting how the one and two stride lines are confused together in single Wi-Fi branch and magnetic field branch performances. The proposed CNN produces well-distinguishable lines for one and two strides. This phenomenon is caused by the fact that the magnetic field branch finds many similar model images as candidates in the whole positioning area. On the other hand, the Wi-Fi branch increases the weights of candidate models that are near to the Wi-Fi branch predictions. As a result, the probability of selecting correct

location improves. As a conclusion, positioning accuracies increase as the lengths of fingerprint images enlarge. However, longer images also bring in larger time delay. Therefore, it is important to find a tradeoff between the accuracy and the delay, and it depends on the application scenario.

D. PERFORMANCES WITH DIFFERENT DEVICE ORIENTATIONS

This experiment tests the performance with different smartphone orientations. A tester holds a smartphone with different device orientations: pointing forward, backward, leftward and rightward. When the smartphone is pointing forward or backward, the smartphone's long edge is parallel to the user moving direction. When the device is leftward or rightward, the long edge is vertical to user direction. The forward mode is the most common scenario for various LBS applications. The e-compass orientation of this mode is in accordance with the user moving direction. Similarly, the backward orientation is inverse to moving direction. Fig. 19 shows the positioning performances of Wi-Fi, magnetic field and united positioning CNN under the abovementioned device orientations. The results indicate that the proposed method performs well under different smartphone orientations. This is because the scheme leverages the RSS of Wi-Fi and the strength of the magnetic field as fingerprint images. The two kinds of features are irrelevant to the different device directions.

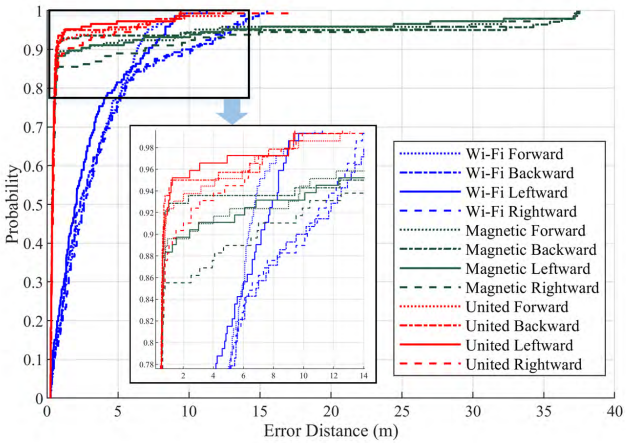


FIGURE 19. Positioning performances with different smartphone orientations.

E. ACCURACIES WITH DIFFERENT USERS

The above experimental results on positioning accuracies are based on the traces mainly collected by one of our authors. In order to examine the system adaptability when used by other people that may walk with different stride lengths and speeds, we recruited other four volunteers (one female and three males) with different heights (between 1.62m to 1.95m) and asked them to test our proposed system in the positioning area. The CNN model is trained based on the data collected by the user one. Fig. 20 denotes the positioning performances

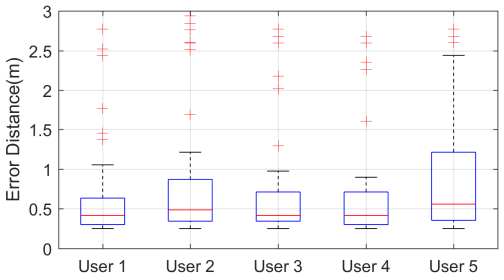


FIGURE 20. Box graph of positioning performance for different users.

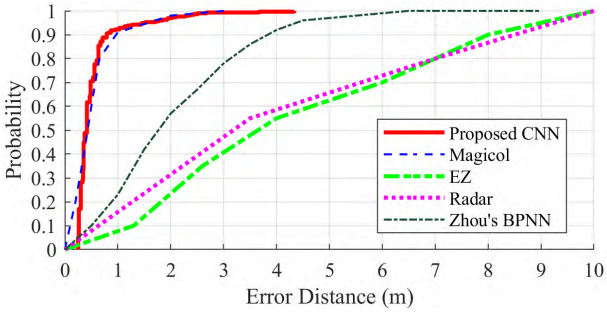


FIGURE 21. Positioning performance comparison with different positioning systems.

using a box graph. Observing the segments between the first and third quartile, we can find that the segment of user one is the shortest and close to the x-axis, indicating that most of the positioning results are near to ground-truth positions. This is because the user one is the model trainer. Compared with user one, the first quartile of the other four users are similar to the user one, but the third quartiles have a larger variation. The first quartile is strongly influenced by the magnetic field branch, so the magnetic field images are stable for different users. On the other hand, Wi-Fi features are prone to be affected by human bodies, therefore the third quartile performances decrease. Although the proposed solution has adopted several techniques to overcome the fluctuation of Wi-Fi signals, it is still difficult to completely eliminate the overfitting problem of Wi-Fi images with a few samples. In the future, we will try crowdsourcing techniques to collect more data as described in [42] and [43] with less human resources and maintaining the Wi-Fi fingerprint updating on time.

F. PERFORMANCE COMPARISON

To evaluate the performance of our proposed algorithm, we compare the localization accuracy with several state-of-the-art positioning systems, i.e., Radar [6], EZ [44], Zhou and Wieser's BPNN method [33] and Magicol [12]. As Fig. 21 reveals, the Radar presents a radio-frequency (RF) based system which records and processes signal strength information of multiple base stations. The EZ system aims at performing indoor localization with no pre-deployment effort. It utilizes opportunistically obtained GPS location

fixes to estimate AP positions and then uses trilateration to calculate user locations with the estimated AP positions. Radar, EZ and Zhou's method only utilize Wi-Fi signals to implement infrastructure-free indoor positioning and achieve decent positioning accuracies. Magicol incorporates Wi-Fi signals, magnetic field signals and dead reckoning with a particle filter, therefore improves positioning performances compared to Wi-Fi only based schemes. Compared with the Magicol method, the proposed fingerprint image method achieves similar accuracies, but the proposed scheme does not rely on dead reckoning, therefore, has no restrictions due to different smartphone orientations.

X. CONCLUSION

This paper proposes a mixed fingerprint image based indoor positioning system utilizing the deep learning. It introduces a novel fingerprint representation scheme to convert Wi-Fi and magnetic field signals into fingerprint images. These high-resolution images are the input of the proposed positioning system and assure the achieving of high positioning accuracies. Compared with traditional positioning methods that needs much effort in designing algorithms for fitting heterogeneous signals, filtering results and tuning parameters, our system uses CNN to automatically learn the mapping between ground-truth positions and fingerprint images. Because of the powerful learning ability of CNN and the orientation insensitivity of the fingerprint image, the proposed positioning scheme achieves orientation-free positioning ability with remarkable accuracies. To make clear the principles of our proposal, we also design and test our system considering Wi-Fi branch only and magnetic field branch only. In order to utilize deep learning technique inside of the indoor positioning context, we also design an indoor position labeling method and an image construction method for creating image-based positioning fingerprints. And the early-stop assisting training method helps the system to learn location features in small training datasets, which makes it applicable in large application. Finally, our experiments convincingly reveal that the proposed positioning method reaches remarkable performances with various smartphone orientations, users, and use cases. For future work, we plan to develop schemes to artificially expanding the training data from sparse Wi-Fi samples to reduce sampling workload.

REFERENCES

[1] G. Zhang and S. V. Rao, "Position localization with impulse ultra wide band," in *Proc. IEEE/ACES Int. Conf. Wireless Commun. Appl. Comput. Electromagn.*, Apr. 2005, pp. 17–22.

[2] Y. Guo, L. Yang, B. Li, T. Liu, and Y. Liu, "RollCaller: User-friendly indoor navigation system using human-item spatial relation," in *Proc. IEEE INFOCOM IEEE Conf. Comput. Commun.*, Apr./May 2014, pp. 2840–2848.

[3] K. Liu, X. Liu, L. Xie, and X. Li, "Towards accurate acoustic localization on a smartphone," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 495–499.

[4] K. Wu, X. Jiang, Y. Youwen, G. Min, and L. M. Ni, "FILA: Fine-grained indoor localization," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 2210–2218.

[5] S. Yang, P. Dessai, M. Verma, and M. Gerla, "FreeLoc: Calibration-free crowdsourced indoor localization," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 2481–2489.

[6] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. 19th Annu. Joint Conf. IEEE Comput. Commun. Soc. (INFOCOM)*, vol. 2, Mar. 2000, pp. 775–784.

[7] N. Hernández, M. Ocaña, J. M. Alonso, and E. Kim, "Continuous space estimation: Increasing WiFi-based indoor localization resolution without increasing the site-survey effort," (in English), *Sensors*, vol. 17, no. 1, Jan. 2017, Art. no. 147.

[8] F. Zhao, H. Luo, X. Zhao, Z. Pang, and H. Park, "HYFI: Hybrid floor identification based on wireless fingerprinting and barometric pressure," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 330–341, Feb. 2017.

[9] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 763–776, Jan. 2017.

[10] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, and L. M. Ni, "CSI-based indoor localization," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 7, pp. 1300–1309, Jul. 2013.

[11] 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, Sep. 2013, Art. no. 73.

[12] Y. Shu, C. Bo, G. Shen, C. Zhao, L. Li, and F. Zhao, "Magicol: Indoor localization using pervasive magnetic field and opportunistic WiFi sensing," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1443–1457, Jul. 2015.

[13] H. Xie, T. Gu, X. Tao, H. Ye, and J. Lv, "MaLoc: A practical magnetic fingerprinting approach to indoor localization using smartphones," presented at the UBICOMP, Seattle, WA, USA, Sep. 2014.

[14] H. Luo, F. Zhao, M. Jiang, H. Ma, and Y. Zhang, "Constructing an indoor floor plan using crowdsourcing based on magnetic fingerprinting," *Sensors*, vol. 17, no. 11, p. 2678, 2017.

[15] W. Shao, H. Luo, F. Zhao, C. Wang, A. Crivello, and T. M. Zahid, "DePos: Accurate orientation-free indoor positioning with deep convolutional neural networks," presented at the Conf. UPINLBS, Wuhan, China, Mar. 2018.

[16] F. Potortì *et al.*, "Comparing the performance of indoor localization systems through the EvAAL framework," *Sensors*, vol. 17, no. 10, p. 2327, 2017.

[17] A. W. S. Au *et al.*, "Indoor tracking and navigation using received signal strength and compressive sensing on a mobile device," *IEEE Trans. Mobile Comput.*, vol. 12, no. 10, pp. 2050–2062, Oct. 2013.

[18] Y. Jiang *et al.*, "ARIEL: Automatic Wi-Fi based room fingerprinting for indoor localization," presented at the ACM Conf. Ubiquitous Comput., Pittsburgh, PA, USA, 2012.

[19] S. He and S.-H. G. Chan, "Sectjunction: Wi-Fi indoor localization based on junction of signal sectors," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2014, pp. 2605–2610.

[20] M. Youssef and A. Agrawala, "The Horus WLAN location determination system," presented at the 3rd Int. Conf. Mobile Syst., Appl., Services, Seattle, WA, USA, 2005.

[21] R. Nandakumar, K. K. Chintalapudi, and V. N. Padmanabhan, "Centaur: Locating devices in an office environment," presented at the 18th Annu. Int. Conf. Mobile Comput. Netw., Istanbul, Turkey, 2012.

[22] R. W. Ouyang, A. K.-S. Wong, C.-T. Lea, and M. Chiang, "Indoor location estimation with reduced calibration exploiting unlabeled data via hybrid generative/discriminative learning," *IEEE Trans. Mobile Comput.*, vol. 11, no. 11, pp. 1613–1626, Nov. 2012.

[23] Z. Xiao, H. Wen, A. Markham, and N. Trigoni, "Lightweight map matching for indoor localisation using conditional random fields," presented at the 13th Int. Symp. Inf. Process. Sensor Netw., Berlin, Germany, 2014.

[24] S. He and K. G. Shin, "Geomagnetism for smartphone-based indoor localization: Challenges, advances, and comparisons," *ACM Comput. Surv.*, vol. 50, no. 6, pp. 1–37, 2017.

[25] H. Wang, S. Sen, A. Elgohary, M. Farid, M. Youssef, and R. R. Choudhury, "No need to war-drive: Unsupervised indoor localization," presented at the 10th Int. Conf. Mobile Syst., Appl., Services, Lake District, U.K., 2012.

[26] Y. Shu, K. G. Shin, T. He, and J. Chen, "Last-mile navigation using smartphones," presented at the 21st Annu. Int. Conf. Mobile Comput. Netw., Paris, France, 2015.

[27] W. Shao *et al.*, "Location fingerprint extraction for magnetic field magnitude based indoor positioning," *J. Sensors*, vol. 2016, Oct. 2016, Art. no. 1945695.

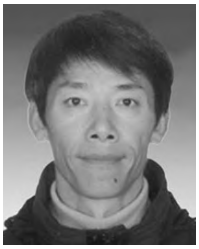
[28] A. K. T. R. Kumar, B. Schäufele, D. Becker, O. Sawade, and I. Radusch, "Indoor localization of vehicles using Deep Learning," in *Proc. IEEE 17th Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Jun. 2016, pp. 1–6.

[29] X. Wang, X. Wang, and S. Mao, "CiFi: Deep convolutional neural networks for indoor localization with 5 GHz Wi-Fi," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.
[30] J. Vieira, E. Leitinger, M. Sarajlic, X. Li, and F. Tufvesson, "Deep convolutional neural networks for massive MIMO fingerprint-based positioning," in *Proc. Pers., Indoor Mobile Radio Commun.*, Oct. 2017, pp. 1–6.
[31] Z. Liu, L. Zhang, Q. Liu, Y. Yin, L. Cheng, and R. Zimmermann, "Fusion of magnetic and visual sensors for indoor localization: Infrastructure-free and more effective," *IEEE Trans. Multimedia*, vol. 19, no. 4, pp. 874–888, Apr. 2017.
[32] M. Nowicki and J. Wietrzykowski, "Low-effort place recognition with WiFi fingerprints using deep learning," in *Proc. Int. Conf. Automat.*, 2017, pp. 575–584.
[33] C. Zhou and A. Wieser, "Application of backpropagation neural networks to both stages of fingerprinting based WIPS," in *Proc. 4th Int. Conf. Ubiquitous Positioning, Indoor Navigat. Location Based Services (UPINLBS)*, 2016, pp. 207–217.
[34] A. Vedaldi and K. Lenc, "MatConvNet: Convolutional neural networks for MATLAB," in *Proc. 23rd ACM Int. Conf. Multimedia*, 2015, pp. 689–692.
[35] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," *Oper. Syst. Design Implement.*, 2016, pp. 265–283.
[36] Y. Jia *et al.*, "Caffe: Convolutional architecture for fast feature embedding," presented at the 22nd ACM Int. Conf. Multimedia, Orlando, FL, USA, 2014, pp. 675–678.
[37] M. Angermann, M. Frassl, M. Doniec, B. J. Julian, and P. Robertson, "Characterization of the indoor magnetic field for applications in localization and mapping," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Nov. 2012, pp. 1–9.
[38] E. Le Grand and S. Thrun, "3-Axis magnetic field mapping and fusion for indoor localization," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst.*, Sep. 2012, pp. 358–364.
[39] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," presented at the 21st Annu. Int. Conf. Mobile Comput. Netw., Paris, France, 2015.
[40] M. A. Youssef, A. Agrawala, and A. U. Shankar, "WLAN location determination via clustering and probability distributions," in *Proc. 1st IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Mar. 2003, pp. 143–150.
[41] W. Shao, H. Luo, F. Zhao, C. Wang, A. Crivello, and M. Z. Tunio, "DePedo: Anti periodic negative-step movement pedometer with deep convolutional neural networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.
[42] W. Zhao, S. Han, R. Hu, W. Meng, and Z. Jia, "Crowdsourcing and multisource fusion-based fingerprint sensing in smartphone localization," *IEEE Sensors J.*, vol. 18, no. 8, pp. 3236–3247, Apr. 2018.
[43] K. Chen, C. Wang, Z. Yin, H. Jiang, and G. Tan, "Slide: Towards fast and accurate mobile fingerprinting for Wi-Fi indoor positioning systems," *IEEE Sensors J.*, vol. 18, no. 3, pp. 1213–1223, Feb. 2018.
[44] K. Chintalapudi, A. P. Iyer, and V. N. Padmanabhan, "Indoor localization without the pain," presented at the 16th Annu. Int. Conf. Mobile Comput. Netw., Chicago, IL, USA, 2010.
[45] F. Potortì, A. Crivello, M. Girolami, P. Barsocchi, and E. Traficante, "Localising crowds through Wi-Fi probes," *Ad Hoc Netw.*, vols. 75–76, pp. 87–97, Jun. 2018.
[46] W. Shao, H. Luo, F. Zhao, and A. Crivello, "Toward improving indoor magnetic field–based positioning system using pedestrian motion models," *Int. J. Distrib. Sensor Netw.*, vol. 14, no. 9, 2018.
[47] W. Shao, H. Luo, F. Zhao, C. Wang, A. Crivello, and M. Z. Tunio, "Mass-centered weight update scheme for particle filter based indoor pedestrian positioning," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–6.

WENHUA SHAO (S'16) is currently pursuing the Ph.D. degree in software engineering with the School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing, China. He is also a Research Assistant with the Institute of Computing Technology, Chinese Academy of Science, Beijing. His current main interests include location-based services, indoor localization, inertial navigation, pervasive computing, neural networks, and machine learning.

HAIYONG LUO is currently an Associate Professor with the Institute of Computing Technology, Chinese Academy of Science, Beijing, China. His main research interests are location-based services, pervasive computing, mobile computing, and Internet of Things.

FANG ZHAO is currently a Professor with the School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing, China. Her research directions are key technology of things, short-range wireless positioning, innovative mobile Internet applications, and mobile information.



YAN MA (M'03) received the degree in computer and telecommunications from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1982. He is currently a Professor with the Institute of Network Technology, BUPT. He has participated in several national key technology research projects, including network management technologies, distance learning systems and multimedia transmissions, IPv6, and embedded systems. He is responsible for the planning, construction, and operation of computer campus network with BUPT. Since 1994, he has been actively participated in the construction and operation of China Education and Research Network (CERNET). He is a member of the expert committee of CERNET, where he is also the Director of the North China. He has participated in the CNGI-CERNET2 network design, construction, and operation, which is one of the backbones of the China Next Generation Internet Project. He is a Senior Member of the China Institute of Communications. He was a recipient of several national science and technology awards from 2007 to 2009.

ZHONGLIANG ZHAO received the B.S. degree from Southeast University, China, the master's degree from the Politecnico di Torino, Italy, and the Ph.D. degree from the University of Bern in 2014, under the supervision of Prof. T. Braun from the Communication and Distributed Systems Group. He is currently a Post-Doctoral Senior Researcher with the University of Bern. His current research interests include mobile ad hoc and sensor networks, mobile cloud computing, SDN/NFV, and urban computing.

ANTONINO CRIVELLO (S'18) received the Ph.D. degree in information engineering and science from the University of Siena, Italy, in 2018. He is currently a Researcher with the Information Science and Technology Institute, Consiglio Nazionale delle Ricerche, Pisa, Italy. His research interests include active and assisted living, indoor localisation, sleep monitoring, and wireless ad hoc networks.