CNN based Indoor Localization using RSS Time-Series

Conference Paper · June 2018

DOI: 10.1109/ISCC.2018.8538530

3 authors:

Mai Ibrahim
King Khalid University

Marwan Torki
Alexandria University

Mustafa Y. Elnainay
Alexandria University

CNN based Indoor Localization using RSS Time-Series

Mai Ibrahim, Marwan Torki and Mustafa ElNainay
Computer and Systems Engineering Department
Alexandria University
Alexandria, Egypt
Email: {eng-mai.ibrahim, mtorki, ymustafa}@alexu.edu.eg

***Abstract*—Indoor localization of mobile nodes is receiving great interest due to the recent advances in mobile devices and the increasing number of location-based services. Fingerprinting based on Wifi received signal strength (RSS) is widely used for indoor localization due to its simplicity and low hardware requirements. However, its positioning accuracy is significantly affected by random fluctuations of RSS values caused by fading and multi-path phenomena. This paper presents a convolutional neural network (CNN) based approach for indoor localization using RSS time-series from wireless local area network (WLAN) access points. Applying CNN on a time-series of RSS readings is expected to reduce the noise and randomness present in separate RSS values and hence improve the localization accuracy. The proposed model is implemented and evaluated on a multi-building and multi-floor dataset, UJIIndoorLoc dataset. The proposed approach provides 100% accuracy for building prediction, 100% accuracy for floor prediction and the mean error in coordinates estimation is 2.77 m.**

***Index Terms*—Received Signal Strength (RSS), fingerprinting, indoor localization, deep learning, convolutional neural network**

I. Introduction

Indoor localization systems are used to determine and track a mobile node location in a closed indoor environment. Recently, a wide range of location-aware services have emerged and created a need for accurate indoor localization systems. These services include disaster management, surveillance, objects tracking in warehouses and targeted advertising based on precise location [1].

Global Positioning System (GPS) is commonly used for outdoor localization but it performs poorly when used in indoor environments. This is because it demands a direct line of sight between the satellites and the receiver to obtain accurate positioning [2]. Furthermore, GPS may consume too much energy to be useful for many applications and different techniques have been proposed to make it more energy efficient [3] [4].

The existing indoor positioning solutions achieve good results based on different measurements from the surrounding environment such as Angle of Arrival, Time of Flight, Return Time of Flight, Received Signal Strength (RSS) and Channel State Information (CSI) [1]. Different techniques are applied on these measurements to determine the node location including fingerprinting [5], lateration [6], dead reckoning [7] and combine these measurements with camera input [8].

This work is partially funded by the Egyptian National Telecommunications Regulatory Authority.

Fingerprinting is one of the most popular schemes in indoor localization. In the offline stage, the localization system builds a database of thorough measurements from reference locations in the target area. Then, in the online stage, the system infers the real-time location by comparing the new measurements with the ones stored in the database. Many existing indoor fingerprinting systems exploit WiFi RSS values (from WLAN access points) as fingerprints due to its simplicity and low hardware requirements [9] [10] [11]. Other systems use CSI data to estimate node location [12] [13]. Unlike RSS, CSI requires modifying the device driver to be able to obtain it from some advanced WiFi network interface cards.

Different machine learning methods are used with RSS fingerprints to build localization models like k-Nearest Neighbors (KNN) [14], Neural Networks [15] and Support Vector Machines (SVM) [15]. However, these methods can suffer from limitations on their ability to fully benefit from the training data to learn complex features. On the other hand, deep learning has been widely applied in various fields where it resulted in great improvement in performance compared to these traditional shallow methods. That is why recent works on RSS-based indoor localization started using deep learning models to estimate nodes locations [16] [17] [18] [19].

These deep learning based models may still provide low positioning accuracy because RSS values correlation with node location can be significantly affected by shadowing and fading. In our proposed solution we tackled this problem by exploiting a time-series of RSS readings from the WLAN access points to estimate node location. Using multiple consecutive RSS readings is expected to remove the noise in separate readings and enhance the localization performance. Additionally, we used convolutional neural network (CNN) to build our localization models to further leverage the temporal dependency between the time-series RSS readings. We used our new approach to address the problem of indoor localization in multi-building and multi-floor environments.

This paper is organized as follows. In Section II, we discuss the related work of indoor localization. We provide a brief background about CNNs in Section III. In Section IV, we describe our proposed solution. We explain the model implementation details and discuss the evaluation results in

Section V. Finally, we provide our conclusions and future work in Section VI.

II. Related Work

Fingerprinting is one of the most important techniques in indoor localization. It is used in different research projects that achieved good positioning accuracy with the aid of machine learning methods that extract and store the main features of survey data in their models. Moreover, it is possible to use fingerprinting with different measurements (fingerprints) including RSS from WLAN access points and CSI data. However, one of the major challenges facing methods that rely on WiFi RSS is the random fluctuations in RSS values. These fluctuations are usually a result of shadowing, fading and multi-path effects.

Some approaches avoid these disadvantages of using RSS fingerprints by exploiting CSI data instead. In [13], Chen et al. present ConFi, a CNN based indoor localization system using CSI that formulates the node localization as a classification problem with labels representing reference points. It builds a CSI features image from the collected CSI data to be the input to CNN then the node location is computed as the weighted centroid of the reference points with high output probabilities. CSI based solutions require extra hardware modifications in the mobile devices to capture the CSI data which makes it harder to apply compared to RSS-based methods that do not need new hardware requirements. Moreover, solving the localization problem as a classification problem based on reference points causes the model accuracy to be sensitive to the number of reference points and their locations. That is why in our proposed solution we use regression to estimate the mobile node location coordinates.

Other solutions attempt to manipulate the RSS data to remove its noise before using it for location estimation. Fang and Lin in [10] followed this approach by using multiple discriminant analysis to extract useful discriminating information from the RSS data. The extracted information is then introduced to a neural network to estimate the node location coordinates. Although the extracted information is expected to be more discriminating and less noisy than the raw RSS values, it can still provide low localization accuracy since it is obtained from a noisy single RSS readings vector. Also, the developed model uses a shallow neural network to predict the node location which may limit its ability to exploit all the important information in the training data. Our proposed solution differs in that it utilizes multiple consecutive RSS readings to predict the mobile node location using a deep neural network which is expected to eliminate or reduce the noise in the collected RSS data and result in better performance.

Deep learning models have been applied on different problems (including computer vision, vehicle routing, natural language processing and speech recognition) where they achieved remarkable improvement. The ability of these models to extract useful information and learn discriminating features from the training data made them attractive to be applied to improve indoor localization performance based on WiFi RSS fingerprints.

Zhang et al. in [16] developed a 4-layer deep neural network (DNN) that generates coarse positioning estimate, which is then refined to produce a final position estimate by a hidden Markov model (HMM) fine localizer. While work in [17] provides an investigation of applying deep belief networks (DBNs) with two different types of Restricted Boltzmann Machines for indoor localization. In both cases, the authors utilize a single vector of RSS data and they focus only on node localization in a single plane. They do not consider the hierarchical nature of multi-building and multi-floor indoor localization that we address in this paper.

Multiple recent works tackled the problem of multi-building and multi-floor indoor localization. The following two papers utilized UJIIndoorLoc dataset [20], the same dataset we use in this paper, for model training and evaluation. In [18] the localization model is based on a single DNN consisting of a stacked autoencoder (SAE) and a feedforward multi-class classifier used for building/floor classification. Their solution is based on a single RSS vector and it does not provide an estimation for the node location coordinates (longitude and latitude). On the other hand, a scalable architecture is proposed in [19] that predicts the node building, floor and location coordinates using a single DNN. Although this approach achieves 99.82 % building hit rate, 91.27 % floor hit rate and 9.29 m position coordinates (longitude, latitude) error, it is still based on a single RSS readings vector to estimate node location, whereas our proposed model uses consecutive time-separated RSS readings to improve the localization accuracy.

III. Preliminaries

In this section, we present background knowledge of convolutional neural networks; their general structure and the applications where they are best suited.

Convolutional neural networks (CNNs) [21] are a special kind of neural networks for processing data that has a known grid-like topology such as time-series data and image data. The general structure of a CNN consists of one or more convolutional layers followed by one or more fully connected layers as in a standard multi-layer neural network as illustrated in Figure 1 [21]. This constrained architecture allows the CNN to leverage the spatial and temporal structure of the input and allows the network to learn more complex features from different parts of the input.

Typically, a convolutional layer consists of three stages. In the first stage, the layer applies a kernel to its input by performing several convolutions in parallel to produce a set of linear activations. In the second stage, each linear activation is run through a nonlinear activation function, such as the rectified linear activation function. In the third stage, a pooling function is used to further modify the output of the layer. Pooling functions replace the output of the net at a certain location with a summary statistic of the nearby outputs. For example, the max pooling operation reports the maximum output within a rectangular neighborhood. Other

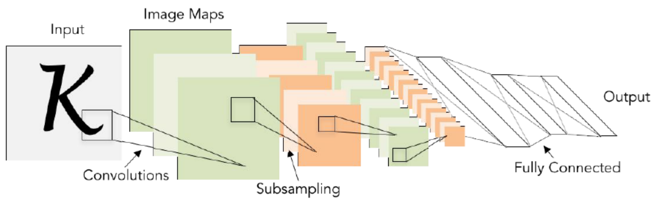

Fig. 1. The general structure of convolutional neural network applied for character recognition [21].

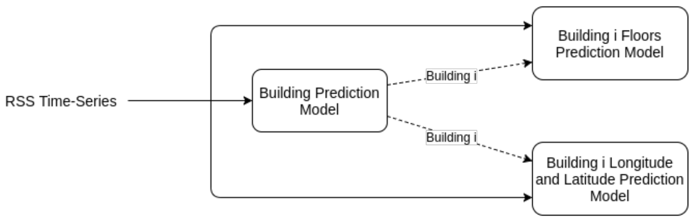

Fig. 2. System Architecture for multi-building and multi-floor indoor localization based on RSS time-series

popular pooling functions include the average of a rectangular neighborhood or a weighted average based on the distance from the central pixel [22].

Although CNNs were originally invented for computer vision, they have been widely applied in different fields where they provided remarkable performance. CNNs were also applied for indoor localization based on CSI fingerprints as proposed in [13]. In this paper, we utilize CNN to build localization models based on an RSS features image generated from a time-series of RSS readings captured by the mobile node from WLAN access points.

IV. Proposed Indoor Localization Approach

The proposed indoor localization scheme addresses mobile node positioning problem in a multi-building and multi-floor environment using a time-series of RSS readings from WLAN access points. Deep learning models are used to predict the mobile node location based on the RSS time series. The node location consists of three parts: its building, floor and location coordinates (longitude and latitude).

The mobile node (i.e., mobile phones, tablets, laptops, etc.) should periodically collect RSS readings from the access points in its WLAN (assume there are N access points). Then, to obtain its location details, the node should send the last T RSS readings to localization server, where the trained models are stored, to estimate the node location based on its RSS time-series readings.

We follow sequential or hierarchical architecture to predict the node location as shown in Figure 2. We start by predicting the node building. Then, based on the predicted building, we use the model tailored specifically to predict this building floors to obtain the node floor. Finally, we use a third model, also based on the node building, to estimate the node location coordinates (longitude and latitude).

Moreover, we present three different approaches of using the time-series of RSS readings to estimate node location. Given T RSS vectors each with N RSS values from the N WLAN access points, we have three options to aggregate these vectors.

- The first approach is to average the T vectors to get a single vector of N values, one from each access point. Then, we input this vector to feedforward DNNs to get the node location. This approach has the drawback that it may lose some of the important information in the multiple RSS values by averaging them.
- Another approach is to concatenate the T readings to get a single RSS vector of TN values, then input this vector to feedforward DNNs to predict the node location. This way is expected to yield better results compared to averaging, but it does not take advantage of the temporal dependency between the T RSS readings in the time-series.
- The third approach is to build RSS features image of size $N \times T$ and use it as the input to CNNs to estimate the node location. The image is built by putting the vectors of the T readings by each other as shown below.

$$\begin{bmatrix} RSS_{1,1} & RSS_{1,2} & \dots & RSS_{1,T} \\ RSS_{2,1} & RSS_{2,2} & \dots & RSS_{2,T} \\ \vdots & \vdots & \vdots & \vdots \\ RSS_{N,1} & RSS_{N,2} & \dots & RSS_{N,T} \end{bmatrix}$$

This method is expected to give the highest accuracy since CNNs leverage the temporal dependency between the RSS vectors that build the features image structure.

V. Implementation and Evaluation Results

We trained and evaluated the proposed solution models using a multi-building and multi-floor dataset, UJIIndoorLoc dataset [20]. Using the same dataset, we implemented localization models based on single RSS readings vector to compare their accuracy against the proposed ones. In this section, we begin by describing the dataset and how we used it. Next, we explain the details of the models we implemented and finally discuss the evaluation results.

A. Dataset Preprocessing

We used UJIIndoorLoc dataset, a public accessible dataset, to build and evaluate our proposed solution. UJIIndoorLoc dataset is a multi-building multi-floor indoor localization database to build indoor positioning systems that rely on WiFi RSS fingerprints. The UJIIndoorLoc data covers three buildingsof Universitat Jaume I, with two buildings having four floors and one building having five floors. It was created in 2013 using 25 Android devices in a WLAN with 520 ($N = 520$) access points. The database consists of 19937 training records and 1111 testing records. The data has 529 attributes, following we describe the ones we used in our model.

- The RSS values received from the 520 access points are negative integer values from -104 dBm (extremely poor signal) to 0 and +100 is used if the access point is not detected.

- The building id, where the RSS readings were captured, takes integer values from 0 to 2.
- The floor id, where the RSS readings were taken, takes integer values from 0 to 4.
- Longitude and latitude of the location where the RSS readings were taken in meters with UTM from WGS84.
- Unix timestamp of when the RSS capture was taken.

We started by building deep neural network models for predicting a node location (i.e. its building, floor, longitude and latitude) based on a single RSS readings vector from the 520 access points in its environment. In these models, we divided the training records and used 80% of them for training and 20% for validation during models parameters tuning. We then evaluated the models accuracy using the 1111 testing records.

To build the proposed scheme models based on multiple consecutive RSS readings, we had to have multiple readings (RSS time-series) from the same location to train the model and evaluate it. Since UJIIndoorLoc dataset does not have enough number of multiple RSS readings for the same location, we had to manipulate the data to obtain the required training and testing data.[1] We started by drawing a grid of $D \times D$ (m^2) cells on the area covered by the training dataset and assigning each record from the 19937 records to a cell based on its longitude and latitude. For each cell records, we sorted them based on their timestamp and put them in groups of T records where the time difference between the oldest and newest reading in the group is less than or equal to S seconds. Each group of T records can then be used as an RSS time-series sample.

We tried different values for D, T and S and we chose the ones that resulted in a sufficiently large number of samples to be used for model training and testing. Using $D = 3$ m, $T = 10$ records and $S = 60$ seconds gave 5484 records of 10 consecutive readings from the same cell. We computed the cell samples longitude and latitude to be the center of this cell. We used 60% of these records for training, 20% for validation and 20% for testing.

We implemented three different approaches to aggregate the 10 RSS readings. The first approach simply averages the 10 readings to get a single RSS vector of 520 values, one from each access point. The second approach concatenates the 10 readings to get a single RSS vector of 5200 (520×10) values. While in the third approach, we build RSS features image of size 520 $\times$ 10 and use it as the input to CNNs to estimate the node location.

Before using the data for model training or evaluation, we normalized the RSS values using Z-score after replacing +100 values with -110 to indicate a very weak signal, as recommended by work in [18]. It is worth mentioning that any location in any building is only covered by few number of access points and this number is different from location to location. Therefore, the models we provide here do not depend on having RSS readings from exactly 520 access points. Actually, in any given location, 27 access points are detected on average.

B. Models Implementation

We built three levels of models to predict the node building, floor and location coordinates (longitude and latitude). Moreover, for each level we trained four different models. One based on single vector of 520 RSS values from the 520 access points in the WLAN. While the other three models are for the three approaches of aggregating the RSS time-series (averaging, concatenation and building RSS features image).

We used early stopping regularization technique in training all the models. For parameters tuning, we performed grid search to find the best parameters. We trained different models for different parameters values and the ones that yield the best validation set error are then chosen as the best parameters. In the following subsections, we only report the best parameters values.[2] However, the following parameters were set to the same values for all models:

- We used rectified linear (RELU) as the activation function for hidden units.
- For classification models, the output units activation function is softmax and the loss function is negative log-likelihood.
- For regression models, the output units activation function is linear and the loss function is mean squared error.
- We used ADAM optimizer and dropout regularization with 0.20 rate.
- In CNN models, we fixed the pooling function to max pooling with kernel size = 2 $\times$ 2 and stride set to 2.

1) Building Prediction: The first step in node localization is to determine the building where the node is located from the three buildings we have in the dataset. We formulated the building prediction as a classification problem by training models to determine if the mobile node belongs to one of three classes/buildings: 0, 1 or 2. Table I shows the parameters of the four models that gave the highest accuracy on validation set.

2) Building Floor Prediction: The second step in determining the node location is to find in which floor of the building the mobile node exists. We implemented models for floor prediction and location coordinates estimation for building two only. This is because we achieved 100% accuracy in building classification problem and the next steps will be the same for the three buildings. We chose building two (which has five floors) specially because it has the largest number of samples in the dataset. Building two floor prediction problem is modeled as a classification problem with the five labels that represent the building five floors. Table II summarizes the best configurations of hidden layers and their corresponding model accuracy on validation set.

[1]The updated dataset used for RSS time-series experiments will be available online for download at the authors homepages.

[2]Source code for the implemented models will be available upon acceptance.

TABLE I
BUILDINGS PREDICTION MODELS PARAMETERS AND ACCURACY ON VALIDATION SET

Input	Convolutional Layers	Hidden Layers Units	Accuracy
Single RSS Vector	-	8, 8, 8	99.8%
Averaged RSS Time-Series	-	8, 8, 8	100%
Concatenated RSS Time-Series	-	8, 8, 8	100%
RSS Time-Series Image	1 layer with three out channels and 2 $\times$ 2 kernel	8	100%

TABLE II
BUILDING TWO FLOORS PREDICTION MODELS PARAMETERS AND ACCURACY ON VALIDATION SET

Input	Convolutional Layers	Hidden Layers Units	Accuracy
Single RSS Vector	-	128, 128, 128	99.9 %
Averaged RSS Time-Series	-	256, 256, 256	97.9%
Concatenated RSS Time-Series	-	8, 8, 8	100%
RSS Time-Series Image	1 layer with 3 out channels and 2 $\times$ 2 kernel	8, 8	100%

3) Longitude and Latitude Estimation: The third and final step is to determine the longitude and latitude of the mobile node. For this step, we trained the DNNs for regression to estimate node location coordinates as continuous values. Each network has two regression outputs that represent the node longitude and latitude. The different models configurations and their mean localization errors in meters are summarized in Table III. The localization error is computed as the Euclidean distance between the predicted coordinates and the actual ones.

C. Evaluation Results

After tuning the models parameters using the validation sets, we used the models that achieved highest accuracy to evaluate their performance on the testing set. In this section, we report the localization accuracy for both the traditional models based on single RSS vector and the proposed solution based on a time-series of RSS readings. Table IV summarizes the results of both sets of models.

It is clear that the proposed solution highly improves the localization performance compared to existing solutions. The time-series based models achieve 100% building prediction accuracy, 100% accuracy for floor prediction and 2.77 m mean localization error in estimating the node longitude and latitude. They obviously outperforms the single RSS vector-based models which give 100% and 91.42% hit rates for building and floor predictions respectively with 10.25 m mean error in coordinates estimation.

The proposed approach also achieves higher localization accuracy when compared to the reported results of other solutions implemented and evaluated on the same dataset but based on separate RSS vectors. According to the results in [23], best building and floor hit rates achieved for the UJIIndoorLoc dataset are 100% and 94% respectively and the mean error in coordinates estimation is 6.20m. Also, the reported results for the single scalable DNN-based system proposed in [19] are 99.82% for building hit rate, 91.27% for floor prediction and 9.29m mean error in coordinates estimation.

This improvement in localization performance proves that using a time-series of RSS readings helps to remove or decrease the noise usually found in a single RSS reading and make it more robust to estimate a node location. The evaluation results also show that the approach used to aggregate the time-series data affects the model accuracy. We can see that models based on the average of the RSS time-series vectors has higher error rates compared to the other two approaches. This is because averaging the time-series can cause losing some of the important information in it. Furthermore, the average-based models do not benefit a lot from having multiple RSS vectors as they return the time-series back to a single RSS vector. Therefore, we witness improvement in accuracy with concatenation based models since they use the RSS vectors as a whole without compressing them in a single one. This improvement comes at the cost of the need to build larger models with more input units (5200 units) compared to the average based ones (only 520 input units). However, these models do not benefit from the relation between the time-series elements. The RSS image-based models result in the highest localization accuracy in the three tasks. In these models, CNNs managed to learn more complex features and hence achieved better localization accuracy thanks to their special architecture that enables them to leverage the temporal dependency between the time-series data. Although CNN models take more training time than DNN ones, their prediction time is comparable to each other. CNN models take 0.84 milliseconds on average to build the RSS time-series image and predict the complete node location (its building, floor and location coordinates) while the single RSS based DNN models take 0.15 milliseconds on average.

VI. CONCLUSION AND FUTURE WORK

In this paper, we propose a novel indoor localization scheme for multi-building and multi-floor environments based on time-series of RSS readings from access points in a WLAN. The new approach is based on deep learning where different deep neural networks are trained to determine the node location information which consists of its building, floor and location

TABLE III
BUILDING TWO LONGITUDE AND LATITUDE ESTIMATION MODELS PARAMETERS AND MEAN ERROR ON VALIDATION SET

Input	Convolutional Layers	Hidden Layers Units	Mean Error (m)
Single RSS Vector	-	150, 150, 150	4.36
Averaged RSS Time-Series	-	300, 300, 300	4.71
Concatenated RSS Time-Series	-	550, 550, 550	4.02
RSS Time-Series Image	2 layers: the first with 8 out channels and 10 $\times$ 3 kernel then the second layer has 4 out channels and 5 $\times$ 3 kernel	128, 128, 128	2.84

TABLE IV
MODELS EVALUATION RESULTS ON TESTING SET

Input	Building Prediction Accuracy	Floor Prediction Accuracy	Localization Mean Error (m)
Single RSS Vector	100%	91.42%	10.25
Averaged RSS Time-Series	99.91%	96.90%	4.39
Concatenated RSS Time-Series	100%	99.66%	3.87
RSS Time-Series Image	**100%**	**100%**	**2.77**

coordinates (longitude and latitude). We proposed three different ways for aggregating the time-series data. The first two aggregation approaches are to average or concatenate the RSS vectors then introducing the result to feedforward DNN-based models to estimate node location. The third approach is to build an RSS time-series image and feed it to CNN-based models to obtain the node location. In addition to these models, we implemented DNN-based models that rely on single RSS vector to evaluate our proposed solution performance against them. All the models were trained and evaluated using UJI-IndoorLoc dataset. The experimental results show that CNN-based models achieve the highest localization accuracy: 100% for building prediction, 100% for floor prediction and 2.77 m mean localization error which introduces large improvement over traditional methods based on single RSS vector that give 100% and 91.42% accuracy for building and floor prediction respectively and 10.25 m mean localization error. This improvement is mainly because using multiple RSS readings accounts for the noise and randomness in a single RSS reading caused by fading, shadowing and multi-path effects.

For future work, we plan to work on a complete scalable version of our approach where the number of trained models is not a function of the number of buildings in the indoor environment. We also plan to validate the proposed model using more datasets.

REFERENCES

[1] F. Zafari et al., "A Survey of Indoor Localization Systems and Technologies," ArXiv preprint, (2017).

[2] R. Sharaf and A. Noureldin, "Sensor Integration for Satellite-Based Vehicular Navigation Using Neural Networks," in IEEE Trans. Neural Netw., vol. 18, no. 2, pp. 589-594, March 2007.

[3] J. Liu et al., "CO-GPS: Energy Efficient GPS Sensing with Cloud Offloading," in IEEE Trans. Mobile Comput., vol. 15, no. 6, pp. 1348-1361, Jun. 1 2016.

[4] P. Misra et al., "Energy efficient GPS acquisition with Sparse-GPS," IPSN, Berlin, 2014, pp. 155-166.

[5] S. He et al., "Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons," in IEEE Commun. Surveys Tuts., vol. 18, no. 1, pp. 466-490, Firstquarter 2016.

[6] J. Yang and Y. Chen, "Indoor Localization Using Improved RSS-Based Lateration Methods," IEEE GlobeCom, Honolulu, HI, 2009, pp. 1-6.

[7] M. Alzantot and M. Youssef, "UPTIME: Ubiquitous pedestrian tracking using mobile phones," IEEE WCNC, Shanghai, 2012, pp. 3204-3209.

[8] S. Papaioannou et al., "Fusion of Radio and Camera Sensor Data for Accurate Indoor Positioning," IEEE MASS, Philadelphia, PA, 2014, pp. 109-117.

[9] M. Youssef and A. Agrawala, "The Horus WLAN location determination system," in ACM MobiSys, Seattle, WA, USA, Jun. 2005, pp. 205-218.

[10] S. H. Fang and T. N. Lin, "Indoor Location System Based on Discriminant-Adaptive Neural Network in IEEE 802.11 Environments," in IEEE Trans. Neural Netw., vol. 19, no. 11, pp. 1973-1978, Nov. 2008.

[11] W. Njima et al., "Smart probabilistic approach with RSSI fingerprinting for indoor localization," SoftCOM, Split, 2017, pp. 1-6.

[12] X. Wang et al., "CSI Phase Fingerprinting for Indoor Localization With a Deep Learning Approach," in IEEE IoT-J, vol. 3, no. 6, pp. 1113-1123, Dec. 2016.

[13] H. Chen, Y. Zhang et al., "ConFi: Convolutional Neural Networks Based Indoor Wi-Fi Localization Using Channel State Information," in IEEE Access, vol. 5, pp. 18066-18074, 2017.

[14] H. Liu et al., "Survey of Wireless Indoor Positioning Techniques and Systems," in IEEE Trans. Syst., Man, Cybern. {C}, vol. 37, no. 6, pp. 1067-1080, Nov. 2007.

[15] S. Dayekh et al., "Cooperative Localization in Mines Using Fingerprinting and Neural Networks," IEEE WCNC, Sydney, NSW, 2010, pp. 1-6.

[16] W. Zhang et al., "Deep neural networks for wireless localization in indoor and outdoor environments", Neurocomputing, vol. 194, pp. 279-287, Jun. 2016.

[17] G. Flix et al., "A fingerprinting indoor localization algorithm based deep learning," ICUFN, Vienna, 2016, pp. 1006-1011.

[18] M. Nowicki and J. Wietrzykowski, "Low-Effort Place Recognition with WiFi Fingerprints Using Deep Learning," in Szewczyk R., Zieliski C., Kaliczyska M. (eds) ICA 2017. Advances in Intelligent Systems and Computing, vol 550. Springer, Cham

[19] K. S. Kim et al., "A Scalable Deep Neural Network Architecture for Multi-Building and Multi-Floor Indoor Localization Based on Wi-Fi Fingerprinting," ArXiv preprint, 2017.

[20] J. Torres-Sospedra et al., "UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems," IPIN, Busan, 2014, pp. 261-270.

[21] Y. Lecun et al., "Gradient-based learning applied to document recognition," in Proc. of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998.

[22] I. Goodfellow et al., Deep Learning., Cambridge, MA, USA:MIT Press, 2016, pp. 326-366 [online] Available: http://www.deeplearningbook.org.

[23] A. Moreira et al., "Wi-Fi fingerprinting in the real world - RTLS@UM at the EvAAL competition," IPIN, Banff, AB, 2015, pp. 1-10.