IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 18, 2019, accepted July 22, 2019, date of publication August 1, 2019, date of current version August 14, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2932469

Joint Time-Frequency RSSI Features for Convolutional Neural Network-Based Indoor Fingerprinting Localization

BEDIONITA SORO, (Student Member, IEEE), AND CHAEWOO LEE, (Member, IEEE)
Department of Electrical and Computer Engineering, Ajou University, Suwon 16499, South Korea

Corresponding author: Chaewoo Lee (cwlee@ajou.ac.kr)

This work was supported by the National Research Foundation (NRF) of Korea funded by the Minister of Education, Science and Technology (MEST) under Grant 2017R1D1A1B03035229.

ABSTRACT The performance of localization methods based on the receiver signal strength (RSS) is significantly affected by the signal strength indicator's (RSSI) instability. To date, there is no adequate approach which significantly reduces the impact of such an instability on the localization accuracy. Hence, in this paper, we propose a continuous wavelet transform (CWT)-based feature extraction method for convolutional neural network (CNN)-based indoor fingerprinting localization method. The proposed feature extraction method uses the continuous wavelet transform to extract the joint time-frequency representation of each raw RSSI data which provides more discriminative information. The extracted features are used with a CNN model to efficiently predict the closest reference points (RPs). Then, a K-nearest neighbors (KNN) model is used to compute the target location. The proposed feature extraction method can be used with a generic deep neural network model to increase the performance where the computing node is not powerful. The proposed method has been evaluated on different datasets and has achieved good performance compared with other well-known existing methods. The experimental results also demonstrated that the proposed approach reduces the influence of RSSI variation.

INDEX TERMS Convolutional neural network, continuous wavelet transform, RSSI-based fingerprinting localization.

I. INTRODUCTION

Indoor positioning and navigation services are playing an important role in society today and are relevant keys for future smart cities [1]–[3]. Unfortunately, with only the Global Positioning System (GPS) it is difficult to achieve a good localization performance in an indoor environment due to the signal attenuation [4]. Therefore, several studies have been carried out to propose alternative solutions to the GPS for indoor environments [5]. Among these methods, approaches based on the RSS, specifically fingerprinting-based localization method [6] is the focus of broad attention in indoor localization literature. Location algorithms based on fingerprinting approach have an important advantage compared with other localization methods in indoor environment: the simplicity of collecting the RSS measurements by today users' handset and the method does not require special devices or network architecture. RSS based fingerprinting localization methods are pattern matching algorithms in which the RSS is used as the pattern. The basic fingerprinting algorithm consists of two main steps: an offline phase and an online phase. During the offline phase, the signal strength of the available Access Points (APs) are collected at predefined positions called reference points and stored in a database called a radio map database. The online phase represents the evaluation step during which the position of target nodes are estimated based on their measurements which are compared with those in the radio map dataset. The estimated position of a target node is the position of the RP whose measurements in the radio map are the most closer to the node's measurements.

However, due to the RSSI instability, this approach barely gives a good location estimation. Hence, to reduce the effect of such an instability of the RSSI on the localization performance, several approaches have been proposed [7]–[9].

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

Among them, methods based on Artificial Neural Network (ANN) are among the promising approaches due to ANN performance on images pattern matching. ANN has been applied in a lot of studies to efficiently increase the pattern matching capability of fingerprinting method, such as approach proposed in [10]. The main challenge facing by RSSI based fingerprinting localization methods is the unpredictable variation of the RSS. Such a variation affects most pattern matching algorithms specifically ANN based fingerprinting methods which perform worse than usual with the RSSI compared with their performance on image data. In fact, an ANN approach requires a good feature extraction method when using the raw RSSI data as input data. Specifically, the deep neural network approach is not quite good enough with RSSI based fingerprinting localization because it cannot implicitly capture the structure of the data from a raw RSSI data as CNN may do. Indeed, CNN generally achieves better results on image recognition than the generic deep neural network (DNN) does.

Several methods are reported in the literature to address the computational cost introduced by the use of CNN models [11]. Specifically Mobilenet [12] which is one of the most well-known light CNN achieving good performance with low computation cost compared to other well-known CNN models [13], [14]. Mobilenet exploits depth-wise separable convolution [15] and pointwise convolution [16] to produce a light CNN model which has few parameters and low computational cost. This model can be used with low computational capability devices such as smartphones. Point-wise CNN model has been used in several studies to assess images dehazing tasks [17], [18]. Recently, a new approach has been proposed in [19] called EfficientNet which improves the performance of existing light CNN such as Mobilenets. Although there are many studies on low computation CNN, the works for indoor positioning based on signal strength remain limited to the generic DNN or standard CNN due to the small size of the available datasets or the small size of the network models. Nevertheless, methods exploiting standard CNN, generally require two-dimensional feature data.

In fact, CNN can efficiently capture the structure of the data and it works more accurately on two-dimensional data than one-dimensional data. Since it is necessary to transform each independent one-dimensional RSSI data into two-dimensional data if we want to accurately utilize the data structure capturing capability of CNN. A simple way to obtain two dimensional RSSI data is to collect multiple samples per location for training as well as for the testing part. However, due to the RSSI variation, such an approach barely achieves a good result and it is time and memory consuming to collect several measurements per RP and being able to perform real-time position estimation. Another approach is to reshape each row sample of the data into two-dimensional data. But such a method requires an intelligent data restructuring such as APs replacement to provide a good feature representation [20].

Additionally, in a previous work [21] we have applied scattering wavelet transform to fingerprinting localization algorithm where the scattering coefficients were used with a DNN model to achieve a good localization performance. This approach has showed that the wavelet transform framework is capable of efficiently extracting features from fluctuating RSSI.

To address the RSSI fluctuation and reliable feature extraction problems, and be able to exploit the advantages of CNN, we propose a feature extraction method that exploits the continuous wavelet transform framework to map each row sample considered as a time series data into two-dimensional joint time-frequency data. Such a transformation provides highly redundant information which is helpful for analyzing and discovering patterns or hidden information from each row data. By using the time and frequency (or scale) representation of the raw data, it is easier for the classifier to learn the discriminative information. Since such a representation provides a good feature representation at different time-scale levels, the proposed method reduces the burden of collecting several samples per location. Also, the proposed feature extraction method can be used with most existing neural network architectures as well as the above-mentioned light CNN model such as Mobilenets. Nonetheless, in our case we limit the implementation to the standard CNN with at most 5 hidden layers which is faster enough for real-time inference.

The rest of the paper is organized as follows. The next section describes related works. Then section III details the proposed method followed by the performance evaluation in section IV. Finally, we conclude the study in section V.

II. RELATED WORKS

The main problem that degrades the RSSI fingerprinting based localization methods is the unpredictable variation of the RSSI signal in indoor environments. There has been numerous studies to investigate the problem and different approaches were proposed [22], [23] to mitigate the impact of RSSI variation on the localization performance. Since the fingerprinting method is firstly a pattern matching approach, there has been a growing interest in applying techniques that incorporate artificial neural network, as in [24] where the authors combined deep neural network model with a hidden Markov approach to accurately achieve location estimation. In the same vein, in [25], [26] a deep autoencoder is used to extract the discriminative features from the RSSI data for a multi-level building classification using a deep neural network.

As the unpredictable variation of the RSSI is the main cause that degrades the fingerprinting method, Channel State Information (CSI) has been considered as an alternative to RSSI in [27], [28]. The CSI provides more information about the radio signal such as phase and amplitude which can be exploited for localization. Unfortunately, the CSI can be affected by the indoor environment due to multipath propagation and fading. Hence, a good pattern matching method such as DNN is required.

However, most existing methods based on DNN do not always provide satisfactory results. Therefore, there has been a growing interest in applying CNN to fingerprinting method, since CNN is very promising on image pattern matching. In [29], a CNN based fingerprinting method is proposed. In that study, the authors exploited subcarriers' information including CSI to produce the feature data. Unfortunately, this approach is more complex than those based on the RSSI and it requires transmission nodes with multiple antennas.

An RSSI based convolutional approach was discussed in [30] where they exploited a time-series approach to construct two-dimensional feature data. This approach requires multiple samples per location or to combine different location data which results in a reduction of the number of samples in the database. Also, as mentioned in section I, collecting several samples per location can be laborious or memory consuming. A method based on Pearson correlation coefficient to produce 2-dimensional features from RSSI data for a CNN based fingerprinting method is introduced in [31]. Additionally, in [32] a CNN based localization method is proposed in which the authors exploited channel impulse response information in indoor nonline-of-sight conditions to reduce the localization error. In their study, the CNN based approach has outperformed the support vector based approach. In [33] an autoencoder approach is proposed using a convolutional neural network to extract reliable features. The method achieved good performance on simulation data. However, there are few methods using CNN with the Wi-Fi fingerprints due to the difficulties to produce reliable two-dimensional data from one-dimensional data.

Moreover, several studies have been carried out to exploit the joint time-frequency information such as Gabor expansion, short time Fourier transform and wavelet transform [34], [35]. Those studies showed that wavelet transform provide better result for time-frequency representation of signals compared to others. In [36], the authors discussed a superfamily of wavelets where they investigated the wavelet suitability for various applications based on its properties. In [37], the wavelet framework and its applications in signal processing from discrete to continuous wavelet transform have been deeply investigated. The wavelet transform has been useful in many medical applications that involve time-series data such as in [38] where a discrete wavelet transform is used to extract feature data from an electrocardiogram signal for support vector machine. As well in [39] a continuous wavelet transform was proposed to extract reliable features of Surface Electromyography. These features were then used by a DNN model to achieve better predictions. Others applications based on time-frequency representation have been investigated in [40], [41].

The above-mentioned works demonstrated the need and challenge of applying CNN to RSSI based indoor fingerprinting localization. They also showed how the continuous wavelet transform can be useful in pattern classification since it well-known in signal processing for time-frequency analysis.

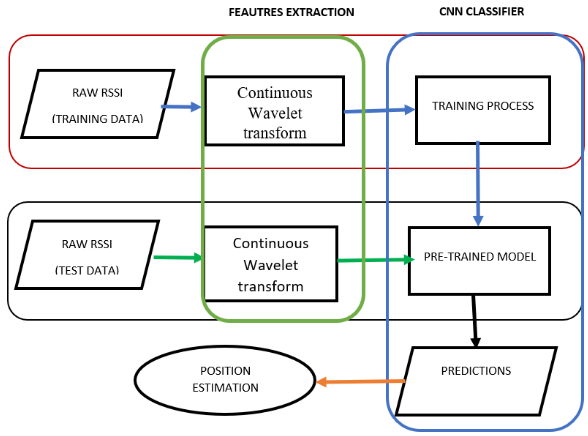

FIGURE 1. Structure of the proposed method.

III. THE PROPOSED SCHEME

This section describes the proposed algorithm. Compared with the previous work [21] where wavelet scattering transform was used with a standard DNN model. In this paper, we apply CWT to extract time-scale features from RSSI samples and use a standard CNN classifier to predict the targets' locations. This approach is designed to exploit time-scale information with CNN using RSSI data. To apply CWT to the RSSI data, we consider each sample as a time-series signal regularly collected at a unit sampling rate and the signal length is the number of APs. Continuous wavelet decomposition is used to extract joint time-frequency information for each row of the RSSI data as presented in the block diagram Fig.1.

The proposed model is presented in Fig.1 and is described as follows: The RSSI data is a numerical array where each row corresponds to a RSSI measurement or RSSI sample from all available APs. Each sample is transformed into a two-dimensional data via continuous wavelet transform operation. This 2D RSSI data is then used to train a convolutional neural network. The output of the convolutional neural network is the closest RPs prediction probability. To estimate the location of the target points, a KNN is applied using these probabilities obtained during the inference time and the RPs' coordinates.

A. CONTINUOUS WAVELET TRANSFORM

Analogously to the Fourier transform, the CWT measures the similarity between a signal and an analyzing function. Fourier transform uses complex exponentials as analyzing functions whereas the CWT uses a wavelet function Ψ. In our study, the analyzing wavelet function is called mother wavelet and is denoted by $\Psi \in L^2(\mathbb{R})$[1] which is a bandwidth and time limited function with zero-mean and unit energy [25]. An example of a standard mother wavelet is a Morlet wavelet where the typical support is $[4, 4]$ as shown in Fig. 2.

[1] denotes the space of square-integrable functions on $\mathbb{R}$.

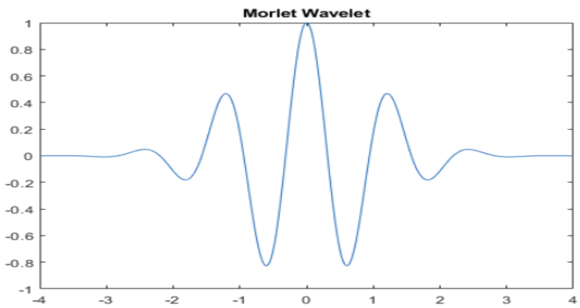

FIGURE 2. Example of mother wavelet.

The boundaries of such wavelet can be extended to $[-8, 8]$. Equation (1) shows the wavelet function shown in Fig. 2.

$$\Psi(t) = e^{\frac{-t^2}{2}} cos(5t). \tag{1}$$

The CWT requires shifted and stretched versions of the analyzing mother wavelet function Ψ which is compared with the signal. We refer to the stretched or compressed versions of the mother wavelet as dilated wavelets in our investigation and denote them by $\Psi_{\alpha,\mu}$. The mother wavelet Ψ is dilated and scaled as defined as follows:

$$\Psi_{\alpha,\mu}(t) = \frac{1}{\sqrt{\alpha}}\Psi(\frac{\mu - t}{\alpha}), \tag{2}$$

where $\mu \in \mathbb{R}$ is a translation factor and $\alpha \in \mathbb{R}^+$ ($\alpha > 0$) is a scaling factor. Let $x(t)$ be the input signal of the wavelet transform at time t. We denote by $U_x(\alpha, \mu)$ the CWT and $< . >$ the inner product. Then we can write the continuous wavelet transform as follows:

$$\begin{aligned} U_x(\alpha, \mu) &= < x(t).\Psi_{\alpha,\mu}(t) > \\ &= \int_{-\infty}^{+\infty} x(t)\frac{1}{\sqrt{\alpha}}\Psi^*(\frac{\mu - t}{\alpha})dt, \end{aligned} \tag{3}$$

where the star exponent denotes the complex conjugate operator. This transformation compares the signal x to the wavelet at various scales α and position μ to produce a function which is a two-dimensional representation of the one-dimensional signal. Obviously, the frequencies can be derived using the scaling factor α. Let ω_Ψ be the angular frequency of the wavelet Ψ. The corresponding scaling frequency $f(\alpha)$ is defined as follows:

$$f(\alpha) = \frac{\omega_\Psi}{2\pi\alpha}. \tag{4}$$

In the paper, we represent each RSSI sample as a signal $X = [X(1), X(2), \ldots X(N)]$ where N is equal to the number of APs that we have considered. The support of the wavelet function $\Psi(t)$ is compact and band limited by definition. We assume that the sampling rate of the signal and the wavelet is 1. Then we can rewrite the above equation (3) with $U_X(\alpha, n)$ representing the continuous wavelet transform of the input vector X as follows:

$$U_X[\alpha, n] = \sum_{k=0}^{N-1} X(k)\frac{1}{\sqrt{\alpha}}\Psi^*(\frac{n - k}{\alpha}), \tag{5}$$

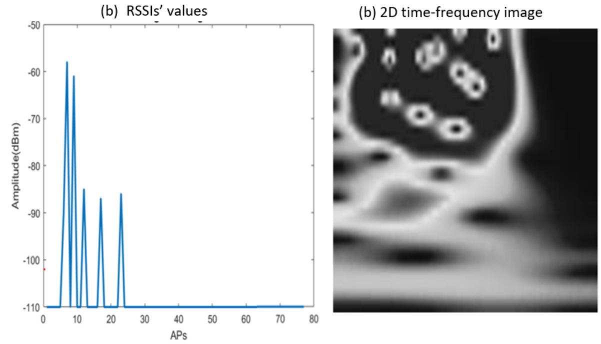

FIGURE 3. Example of RSSI sample and its continuous wavelet transform result.

where $n \in \mathbb{Z}$, and N is the number of APs or the length of each row RSSI data. For a mathematical convenience we define the stretched wavelet as $\Psi_\alpha(n)$ in equation (6).

$$\Psi_\alpha(n) = \frac{1}{\sqrt{\alpha}}\Psi(\frac{n}{\alpha}), \tag{6}$$

where $\Psi_\alpha(n) = 0$ if $n < 0$. Based on the stretched wavelet in equation (6), we can rewrite equation (5) as follows:

$$\begin{aligned} U_X[\alpha, n] &= \sum_{k=0}^{N-1} X(k)\Psi_\alpha^*(n - k) \\ &= X \star \Psi_\alpha^*(n), \end{aligned} \tag{7}$$

where $n = 0, 1, \ldots, N - 1$, and the star ($\star$) denotes the convolutional operator. Such approach of computing the CWT is known as fast continuous wavelet transform.

To generate the two-dimensional features from the one-dimensional data, we define a sequence of M scaling factors $\alpha = \alpha_1, \alpha_2, \ldots, \alpha_M$. Then, we compute the continuous wavelet transform for each scaling factor to produce a $M \times N$ joint time-frequency or time-scale representation feature matrix for each row data. Fig. 3 shows a sample of RSSI measurement from a RP where 77 APs are considered and its CWT computed with 37 scaling factors (frequency components). We can easily change the output image's height of the CWT by changing the number of scaling factors. However, the output image's width is related to the number of APs (input length) which can be large and computationally expensive. Therefore, dimensionality reduction can be applied to reduce the input length to an affordable length and the CWT computation cost. In Fig. 3.a, from around the 25^{th} AP to the 77^{th} AP there was no available RSS signal for these APs, so their RSSI values are set to -110 dBm.

B. ARTIFICIAL NEURAL NETWORKS MODELS

In this section, we describe briefly the convolutional neural network architecture used in our algorithm. To accurately predict the closest RPs from the extracted feature data, we used a CNN model with four hidden layers and one fully connected layer as shown in Fig. 4. A convolutional neural network is a sequence of layers, where each layer transforms one

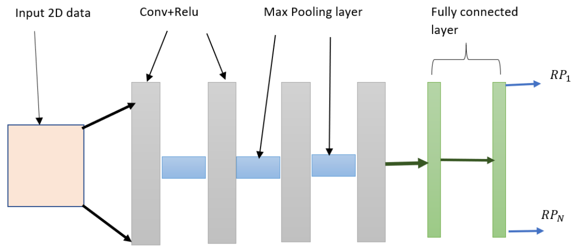

FIGURE 4. Proposed convolutional neural network architecture.

TABLE 1. CNN parameters.

Parameters	Values and expressions
Number of convolution layers	4
hidden layers' activation function	relu
Number of neurons per hidden layer	128
Optimization method	AdamOptimizer
Weights decay	0.0
Dropout	0.2
validation split	0.1
Output layer	SoftMax
filter	2×2
pooling size	2×2
Number of layers for fully connected	1
Number of hidden neurons of the FC	512
Loss function	binary cross entropy
Number of epochs	100
Batch size	16
Learning rate	0.001

volume of activation to another using differentiable functions. A simple CNN consists of three main layers: convolutional layer, pooling layer, and fully-connected layer. The input time-frequency representation of the raw RSSI is assimilated to a grayscale image data where the depth is one.

The convolutional layer performs a convolutional operation on the input using two-dimensional filters to decompose the input image in cluster outputs that are passed through a linear rectifier activation function relu such as $relu(x) = max(0, x)$. Each cluster of neurons is downsampled to a single neuron using the max pooling operation. A max pooling operation uses the maximum value of each cluster of neurons produced from the previous layer. The last part of the CNN is a fully-connected layer (FC) which is a generic neural network classifier. In the FC, the feature image from the last convolutional layer is flattened and passed through different layers where every neuron in one layer is connected to every neuron in the next layer to the output layer. SoftMax activation function [42] is commonly used as output layer for classification. Therefore, in this study, the CNN's output layer consists of a SoftMax layer which outputs the closeness probabilities of different RPs. The probabilities obtained during the online phase are then used by a KNN method to estimate the target node's location. Table 1 describes the CNN's parameters.

To compute the network predictions, let the result of the last hidden layer output be $H_L = [h_1, h_2, \ldots, h_M]^T$ which is a column vector and the output layer's weights be $W_L = [w_{ij}]$, $(i = 1, \ldots, N, j = 1, \ldots, M)$. $W_L H_L = [z_1, z_2, \ldots, z_N]$ represents the input feature vector of the SoftMax layer, where N is the number of RPs. Then, the i^{th} output (p_i) of the neural network or SoftMax output corresponds to the i^{th} RP's prediction probability as follows:

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}}. \tag{8}$$

To estimate the location of a given node T, we consider only the first three RPs whose predictions are greater than the defined threshold denoted by *thres*. Let Γ_T be the first three RPs' index for which the prediction probabilities are greater or equal to the threshold and (x_i, y_i), $i \in \Gamma_T$, the corresponding coordinates. Then we compute the estimated position of the target node (T) defined as $(\hat{x}, \hat{y})$ as follows:

$$\hat{x} = \frac{\sum_{i\in\Gamma_T} x_i.p_i}{\sum_{i\in\Gamma_T} p_i}$$
$$\hat{y} = \frac{\sum_{i\in\Gamma_T} y_i.p_i}{\sum_{i\in\Gamma_T} p_i}, \tag{9}$$

where $p_i \geq thres$.

To describe the localization error evaluation, we assume that there are q target points with real and estimated coordinates respectively as (x_i, y_i) and $(\hat{x}_i, \hat{y}_i)$, $i = 1, \ldots, q$. Then, the localization error is computed as follows:

$$err = \frac{1}{q}\sum_{j=1}^{q}\sqrt{(x_j - \hat{x}_j)^2 + (y_j - \hat{y}_j)^2}. \tag{10}$$

The localization error is evaluated by considering the average error introduced by each target point.

IV. PERFORMANCE EVALUATION

All experiments were conducted using the same desktop with the following characteristics: Intel core i5-2500 64bits CPU 3.30Ghz quad core, Intel Sandybridge graphic, and 8Gb of RAM. Also we only used the CPU to perform all experiments.

To evaluate the performance of the proposed method, we carried out different types of experiments using different RSSI datasets. Then we compared the proposed algorithm performance with those of the neural network model used in [10], [24]–[26] which is a stacked autoencoder classifier model and referred as DNN+SAE or SAE+DNN. We used the SAE+DNN as DNN approach for the performance comparison with the proposed method because it is one of the commonly implemented state-of-the-art DNN based fingerprinting approach. The SAE+DNN model used in this paper, has five hidden layers for the autoencoder whose number of neurons are respectively 128, 64, 32, 64, and 128. The classifier part uses the encoded part of the autoencoder(128-64-32) and two extra hidden layers are added with number of neurons respectively 128 each one. We ran the experiments for different random seed values to demonstrate the stability of each methods.

FIGURE 5. Corridor environment.

TABLE 2. Results of the experiment on the corridor.

Algorithms	Random seed	Average localization error (meter)
Proposed method	0	1.10
	1	1.11
	7	1.17
	42	0.90
	None	0.98
Average		1.05
Standard deviation		0.109
SAE+DNN	0	1.24
	1	1.45
	7	1.83
	42	1.21
	None	1.57
Average		1.46
Standard deviation		0.255
KNN		2.02

A. PERFORMANCE EVALUATION WITH OUR DATA

This experiment was carried out in a corridor area of 50m× 1.95m at Ajou University. Fig. 5 presents the part of the corridor in which the experiment has been carried out.

In the database, there are 100 samples per reference point collected with a frequency of 10 seconds. The dataset has 21 RPs and 11 test points for which only one sample per test point was collected. And only 36 APs RSSI were used in the process of experimentation. Then we defined the number of scaling factors equal to the half of the number of APs and used a Morlet wavelet. The test results of our method compared to other techniques as presented in Table 2 showed that the proposed method outperformed other methods. With the proposed feature extraction method, we achieved good localization performance with the CNN as well as with DNN. This demonstrates that the CWT produces discriminative information that works with different neural network architectures.

The inference times in this experiment for 11 test samples were approximately 20ms for the proposed CNN method and 5ms for the SAE+DNN method. However, the proposed method achieved the best results with lower variance. Thus, it is more stable than the SAE+DNN model.

B. PERFORMANCE EVALUATION WITH A PUBLIC DATASET

In this section, we evaluated the proposed method using a publicly available dataset [43] which consists of 25 subsets of datasets collected at different periods independently from 620 APs. In that database, there are 12 samples collected per location for the training data with 48 RPs. The data were collected in a building at the fifth and third floor only as described in [43]. Therefore, it is suitable for testing the impact of RSSI variation in the same environment and location at different dates. To evaluate the proposed model on that dataset, we first considered the first subset of the database and defined two experimental scenarios. We downsampled the number of APs to 77 before applying the CWT. However, we noticed that the way we downsampled the number of APs was affecting the floor prediction rate of the SAE+DNN in the first scenario. Since, we used the raw RSSI without downsampling and that have significantly improved the SAE+DNN's performance on floor prediction but did not reduce significantly the prediction error. Although the floor hit rate has been improved in the first scenario, it has been degraded in the second experiment and also the inference time has increased. We used 576 test samples for both scenarios.

TABLE 3. Results of the experiment for the first scenario.

Algorithms	Random seed	Floor hit ratio(%)	Average localization error (meter)
Proposed method	0	99.98	0.002
	1	100	0.0
	7	99.82	0.003
	42	100	0.0
	None	100	0.0
Average		99.96	0.001
Standard deviation		0.0787	0.0014
SAE+DNN	0	97.33	5.46
	1	96.73	7.15
	7	98.37	5.06
	42	99.21	4.82
	None	64.68	4.94
Average		98.13	5.48
Standard deviation		1.0763	0.961
KNN		100.0	3.17

The first scenario aimed to check the stability of the proposed method in relation to the RSSI instability. Hence, we only considered the training data ($tr01rss$) and test data ($tst01rss$) which have been collected at the same location. he results of that experiment are presented in Table 3.

We noticed that, in Table 3, the proposed algorithm achieved almost 100% on floor classification and matched most of the test points to their correct location. The result demonstrates that the proposed scheme can significantly reduce the influence of RSSI variation compared with other methods. The proposed method also has a small standard deviation in both floors and position estimation compared to SAE+DNN approach.

Secondly, we evaluated the proposed algorithm using the first subset of the test data of the 25th month and the first subset of the training data of the first month dataset. Since the training and the test data were collected at different time periods, they are likely different due to the RSSI fluctuation. The purpose of this experiment is to evaluate the performance of the proposed method on a long period compared

TABLE 4. Results of the experiment of the second scenario.

Algorithms	Random seed	Floor hit ratio(%)	Average localization error (meter)
Proposed method	0	98.91	4.30
	1	98.88	4.35
	7	99.65	4.62
	42	99.89	3.13
	None	99.80	4.10
Average		99.42	4.10
Standard deviation		0.492	0.573
SAE+DNN	0	98.95	5.12
	1	93.40	6.09
	7	97.22	5.50
	42	96.52	6.18
	None	97.91	5.66
Average		96.80	5.71
Standard deviation		2.10	0.436
KNN		50.0	7.46

with SAE+DNN. Table 4 shows the localization error of the second scenario.

Although, the proposed method achieved similar results on floor classification with the SAE+DNN model, it has achieved better localization accuracy compared with the DNN+SAE and KNN methods. The inference times in both scenarios for 576 samples were approximately 1.5s for the proposed CNN method and 25ms for the SAE+DNN approach. We noticed that in this database the floor misclassification did not affect significantly the localization error when using multilabel classification. This is due to the way the coordinate system and the RPs were defined by the authors of the data set. In fact, they defined the RPs of upper floor to be correspondingly above those in the lower floor.

In this study, we have proposed a CWT-based time-frequency feature extraction for RSSI based fingerprinting localization. CWT has been applied to each RSSI samples considered as a time-series data. Then a deep convolutional neural network model is used to efficiently predict the closest position. The results showed that the proposed method has achieved good performance in all experiments. They have demonstrated that the variance of the localization error for the proposed method is lower compared to other methods.

The present study provides an insight of how to exploit time-frequency or time scaled information of the RSSI. The results have proved that it is a promising approach. Regarding the scalability of the proposed algorithm, a dimension reduction method may be applied to reduced the samples' length in case there are lots of APs. We have also carried out a test using Mobilenet [12] with the proposed feature extraction method. However, this test result was not satisfactory. The Mobilenet model give less than 20% where the proposed CNN model produced 100%. We believe that further investigations are required to apply Mobilenet to RSSI based fingerprinting.

V. CONCLUSION

In this paper, we have proposed a new approach for extracting discriminative features for indoor fingerprinting localization algorithm. The proposed method uses continuous wavelet transform to extract time-frequency or time-scaling information from each RSSI samples. Then we used a CNN classifier to predict the closest RPs. The proposed method achieved a better performance on the overall experiment than some existing deep neural network-based method. Since the feature extraction method exploits time-scaling information of the signal, it can reduce the burden of collecting several samples per location and there is no need of reshaping the dataset to produce 2-dimensional data. Furthermore, the proposed feature extraction method can provide input features to most neural network models. In future works, we hope to extend our investigation to the use of light CNN as Mobilnet and Efficientnet with RSSI for indoor localization.

REFERENCES

[1] M. Mohammadi and A. Al-Fuqaha, "Enabling cognitive smart cities using big data and machine learning: Approaches and challenges," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 94–101, Feb. 2018.

[2] S. Sadowski and P. Spachos, "RSSI-based indoor localization with the Internet of Things," *IEEE Access*, vol. 6, pp. 30149–30161, 2018.

[3] V. D. Bianchi, P. Ciampolini, and I. De Munari, "RSSI-based indoor localization and identification for ZigBee wireless sensor networks in smart homes," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 2, pp. 566–575, Feb. 2019.

[4] G. Dedes and A. G. Dempster, "Indoor GPS positioning—Challenges and opportunities," in *Proc. IEEE 62nd Veh. Tech. Conf. (VTC-Fall)*, vol. 1, Sep. 2005, pp. 412–415.

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

[6] A. Li, J. Fu, A. Yang, and H. Shen, "A new RSS fingerprinting-based location discovery method under sparse reference point conditions," *IEEE Access*, vol. 7, pp. 13945–13959, 2019.

[7] B. Molina, E. Olivares, C. E. Palau, and M. Esteve, "A multimodal fingerprint-based indoor positioning system for airports," *IEEE Access*, vol. 6, pp. 10092–10106, 2018.

[8] Y. Fu, P. Chen, S. Yang, and J. Tang, "An indoor localization algorithm based on continuous feature scaling and outlier deleting," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 1108–1115, Apr. 2018.

[9] X. Zhang, A. K.-S. Wong, C.-T. Lea, and R. S.-K. Cheng , "Unambiguous association of crowd-Sourced radio maps to floor plans for indoor localization," *IEEE Trans. Mobile Comput.*, vol. 17, no. 2, pp. 488–502, Feb. 2018.

[10] B. Soro and C. Lee, "Performance comparison of indoor fingerprinting techniques based on artificial neural network," in *Proc. IEEE Region Conf. TENCON*, Oct. 2018, pp. 28–31.

[11] L. Kaiser, A. N. Gomez, N. Shazeer, A. Vaswani, N. Parmar, L. Jones, and J. Uszkoreit, "One model to learn them all," 2017, *arXiv:1706.05137*. [Online]. Available: https://arxiv.org/abs/1706.05137

[12] Andrew G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*. [Online]. Available: https://arxiv.org/abs/1704.04861

[13] S. Christian, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," 2014, *arXiv:1409.4842*. [Online]. Available: https://arxiv.org/abs/1409.4842

[14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556

[15] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," 2016, *arXiv:1610.02357*. [Online]. Available: https://arxiv.org/abs/1610.02357

[16] B.-S. Hua, M.-K. Tran, and S.-K. Yeung, "Pointwise convolutional neural networks," 2017, *arXiv:1712.05245*. [Online]. Available: https://arxiv.org/abs/1712.05245

[17] J. Zhang, Y. Cao, Y. Wang, C. Wen, and C. W. Chen, "Fully point-wise convolutional neural network for modeling statistical regularities in natural images," in *Proc. ACM Multimedia Conf.*, Oct. 2018, pp. 984–992.

[18] J. Zhang and D. Tao, "FAMED-Net: A fast and accurate multi-scale end-to-end Dehazing network," 2019, *arXiv:1906.04334*. [Online]. Available: https://arxiv.org/abs/1906.04334

[19] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, May 2019, pp. 6105–6114.

[20] Z. Zhong, Z. Tang, X. Li, T. Yuan, Y. Yang, M. Wei, Y. Zhang, R. Sheng, N. Grant, C. Ling, X. Huan, K. S. Kim, and S. Lee, "XJTLUIndoorLoc: A new fingerprinting database for indoor localization and trajectory estimation based on Wi-Fi RSS and geomagnetic field," 2018, *arXiv:1810.07377*. [Online]. Available: https://arxiv.org/abs/1810.07377

[21] B. Soro and C. Lee, "A wavelet scattering feature extraction approach for deep neural network based indoor fingerprinting localization," *Sensors*, vol. 19, no. 8, p. 1790, Apr. 2019.

[22] S. Sahin, H. Ozcan, and K. Kucuk, "Smarttag: An indoor positioning system based on smart transmit power scheme using active tags," *IEEE Access*, vol. 6, pp. 23500–23510, 2018.

[23] K.-H. Chow, H. He, J. Tan, and S.-H. G. Chan, "Efficient locality classification for indoor fingerprint-based systems," *IEEE Trans. Mobile Comput.*, vol. 18, no. 2, pp. 290–304, Feb. 2019.

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

[25] J. Wietrzykowski and M. Nowicki, "Low-effort place recognition with WiFi fingerprints using deep learning," 2017, *arXiv:1611.02049*. [Online]. Available: https://arxiv.org/abs/1611.02049

[26] K. S. Kim, L. Sanghyuk, and K. Huang , "A scalable deep neural network architecture for multi-building and multi-floor indoor localization based on Wi-Fi fingerprinting," *Big Data Anal.*, vol. 3, no. 4, p. 4, Apr. 2018.

[27] X. Wang, L. Gao, S. Mao, and S. Pandey, "DeepFi: Deep learning for indoor fingerprinting using channel state information," *IEEE Wireless Commun. Netw. Conf. (WCNC)*, New Orleans, LA, USA, Mar. 2015, pp. 1666–1671.

[28] Q. Song, S. Guo, X. Liu, and Y. Yang, "CSI amplitude fingerprinting-based NB-IoT indoor localization," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1494–1504, Jun. 2018.

[29] H. Chen, Y. Zhang, W. Li, X. Tao, and P. Zhang, "ConFi: Convolutional neural networks based indoor Wi-Fi localization using channel state information," *IEEE Access*, vol. 5, pp. 18066–18074, 2017.

[30] M. Ibrahim, M. Torki, and M. ElNainay, "CNN based indoor localization using RSS time-series," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jun. 2018, pp. 01044–01049.

[31] A. Mittal, S. Tiku, and S. Pasricha, "Adapting convolutional neural networks for indoor localization with smart mobile devices," in *Proc. Great Lakes Symp. VLSI*, Chicago, IL, USA, May 2018, pp. 117–122.

[32] K. Bregar and M. Mohorčič, "Improving indoor localization using convolutional neural networks on computationally restricted devices," *IEEE Access*, vol. 6, pp. 17429–17441, 2018.

[33] Z. E. Khatab, A. Hajihoseini, and S. A. Ghorashi, "A fingerprint method for indoor localization using autoencoder based deep extreme learning machine," *IEEE Sensors Lett.*, vol. 2, no. 2, Mar. 2018, Art. no. 6000204.

[34] H.-G. Stark, *Wavelets and Signal Processing: An Application-Based Introduction*. Berlin, Germany: Springer, 2005.

[35] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed. New York, NY, USA: Academic, 2008.

[36] J. M. Lilly and S. C. Olhede, "Generalized morse wavelets as a superfamily of analytic wavelets," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 6036–6041, Nov. 2012.

[37] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bull. Amer. Meteorol. Soc.*, vol. 79, no. 1, pp. 61–78, Jan. 1998.

[38] Q. Zhao and L. Zhang, "ECG feature extraction and classification using wavelet transform and support vector machines," in *Proc. IEEE Int. Conf. Neural Netw. Brain (ICNN&B)*, vol. 2, Oct. 2005, pp. 1089–1092.

[39] J. Kilby and H. G. Hossein, "Extracting effective features of SEMG using continuous wavelet transform," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, New York, NY, USA, Aug. 2006, pp. 1704–1707.

[40] J. M. Lilly, "Element analysis: A wavelet-based method for analysing time-localized events in noisy time series," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 473, no. 2200, Apr. 2017, Art. no. 20160776.

[41] V. Pukhova, "Adaptive time-frequency analysis of signals in AFM," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 443, no. 1, Nov. 2018, Art. no. 012026.

[42] I. Goodfellow, Y. Bengio, and A. Courville, "6.2.2.3 softmax units for multinoulli output distributions," in *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016, pp. 180–184.

[43] G. M. Mendoza-Silva, P. Richter, J. Torres-Sospedra, E. S. Lohan, and J. Huerta, "Long-term WiFi fingerprinting dataset for research on robust indoor positioning," *Data*, vol. 3, no. 1, p. 3, Jan. 2018.

BEDIONITA SORO received the electronic engineering degree from the Institut Polytechnique Houphoit Boigny (INP-HB), Yamoussoukro, Côte d'Ivoire, in 2015. He is currently pursuing the master's degree in electrical and computer engineering with Ajou University, Suwon, South Korea. His research interests include the wireless sensor networks and machine learning application. He is currently involving in indoor positioning and navigation based on fingerprinting method.

CHAEWOO LEE received the B.S. degree in control and instrumentation engineering from Seoul National University, in 1985, the M.S. degree in electrical and computer engineering from the Korea Advanced Institute of Science and Technology (KAIST), in 1988, and the Ph.D. degree in electrical and computer engineering from the University of Iowa, in 1995. He was with Korea Telecom and Lucent Technologies, from 1988 to 1999. Since March 2002, he has been with the Department of Electrical and Computer Engineering, Ajou University. His research interests include multimedia communication systems, and the IoT systems and applications.