

A survey of deep learning approaches for WiFi-based indoor positioning

Xu Feng, Khuong An Nguyen & Zhiyuan Luo

To cite this article: Xu Feng, Khuong An Nguyen & Zhiyuan Luo (2021): A survey of deep learning approaches for WiFi-based indoor positioning, Journal of Information and Telecommunication, DOI: [10.1080/24751839.2021.1975425](https://doi.org/10.1080/24751839.2021.1975425)

To link to this article: <https://doi.org/10.1080/24751839.2021.1975425>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 20 Sep 2021.



Submit your article to this journal [↗](#)



Article views: 1472



View related articles [↗](#)



View Crossmark data [↗](#)



A survey of deep learning approaches for WiFi-based indoor positioning

Xu Feng ^a, Khuong An Nguyen ^a and Zhiyuan Luo ^b

^aSchool of Computing, Engineering and Mathematics, University of Brighton, Brighton, UK; ^bDepartment of Computer Science, Royal Holloway University of London, Surrey, UK

One of the most popular approaches for indoor positioning is WiFi fingerprinting, which has been intrinsically tackled as a traditional machine learning problem since the beginning, to achieve a few metres of accuracy on average. In recent years, deep learning has emerged as an alternative approach, with a large number of publications reporting sub-metre positioning accuracy. Therefore, this survey presents a timely, comprehensive review of the most interesting deep learning methods being used for WiFi fingerprinting. In doing so, we aim to identify the most efficient neural networks, under a variety of positioning evaluation metrics for different readers. We will demonstrate that despite the new emerging WiFi signal measures (i.e. CSI and RTT), RSS produces competitive performances under deep learning. We will also show that simple neural networks outperform more complex ones in certain environments.

ARTICLE HISTORY

Received 11 June 2021

Accepted 25 August 2021

KEYWORDS

Deep learning; neural network; WiFi fingerprinting

1. Introduction

In terms of the technologies applied in indoor positioning, one of the most popular methods is WiFi. Thanks to the ubiquity of smartphones, the past decade has witnessed the proliferation of WiFi devices including computers, smartphones, and multitudinous Access Points (APs). Offices, hospitals, shopping malls and factories are now densely populated with WiFi APs to provide the Internet services for the users. Therefore, it is convenient to take advantage of such technology for indoor positioning. Nevertheless, proposing a sub-metre level accuracy WiFi-based navigation system is still a research challenge.

The problem with this type of system is the high-dimensional data. To accurately locate the targeted person or object in the indoor environment, the system needs to analyse the signals from hundreds of nearby WiFi APs. Traditional machine learning methods are slow in dealing with such high-dimensional dataset. However, recent systems leverage these big data by applying deep learning, a relatively new machine learning approach to provide new representation of input data. The nature of deep learning makes it suitable to deal with massive amount of high-dimensional data. Other than its capability of

CONTACT Xu Feng  x.feng@brighton.ac.uk

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

extracting hierarchical information from discrete input data, deep learning could also generate accurate position estimation directly. Similar to traditional machine learning methods, deep learning could be modified as a regressor or a classifier to perform distinguishing positioning tasks. Thus, in current WiFi-based indoor positioning systems, deep learning could be used as a feature extraction method, or, as a positioning prediction method.

The authors have pooled over 1000 research papers on Google Scholar that satisfy the following three conditions:

- Must contain the ‘indoor’ and ‘WiFi’ keywords.
- Contains at least one of the following keywords: ‘navigation’, ‘positioning’, ‘localization’, ‘tracking’.
- Contains at least one of the following keywords: ‘deep learning’, ‘neural network’, ‘CNN’, ‘ANN’, ‘RNN’.

Then, each paper was carefully reviewed to ascertain its relevance and suitability for this article. In the end, more than 150 research papers were chosen for further detailed comparison and analysis. In doing so, we aim to answer the following research questions.

- *What is the most accurate WiFi signal measure for indoor positioning systems?* Received signal strength, channel state information, and round-trip time, are the most popular measures of the WiFi signal reported in the literature.
- *What are the most efficient neural networks for WiFi-based indoor positioning systems?* The common perception was that neural networks with complex structures would deliver better results (e.g. CNN with tens or hundreds of hidden layers), widely reported for image classification. Does the same hypothesis hold for WiFi indoor positioning?

1.1. Review scope

Indoor positioning only focuses on predicting the user’s location in constrained environments (e.g. office buildings, hospitals, train stations, shopping malls, etc.). Such an environment often contains multiple rooms, corridors and floors, and is crowded with furniture, walls and people. Therefore, the electromagnetic signals are usually blocked, attenuated and reflected when travelling in an indoor environment. All systems covered in this review perform their experiments in the indoor environments under the above limited conditions.

Although WiFi indoor positioning and deep learning have been around independently, it is not until recent years that researchers started to apply deep learning and neural networks into WiFi indoor positioning. Since the beginning, traditional machine learning methods (i.e. shallow learning) have had a virtual monopoly in this area. However, shallow learning was not able to effectively make use of the massive amount of high-dimensional data, and could hardly reach sub-metre level accuracy. The advance of deep learning enables the researchers to find valid representations of the WiFi data. This trend motivates the composition of this review.

There have been several surveys in the intersection of deep learning, neural networks and indoor positioning. However, most of them either focus on all related machine

learning approaches, or cover a wide range of indoor technologies. In contrast, this review emphasizes on just WiFi-based systems using deep learning and neural networks to provide the readers with a concise understanding of this emerging approach.

1.2. Article's contributions

The contributions of this review article are as follows.

- We divide deep learning-based systems into two categories: those using deep learning as a feature extraction method and those employing it as a positioning prediction method. All comparisons are made individually within each category.
- We particularly analyse the effect of different WiFi signal measures as the input of deep learning-based systems.
- We thoroughly consider the results from various systems with different types of neural networks to find out the most accurate solution for WiFi indoor positioning systems.
- We derive a standard set of evaluation metrics to assess and compare the performance of over 150 deep learning-based indoor positioning systems.

The remaining of the article is organized as follows. Section 2 is concerned with the basic idea of WiFi fingerprinting. Section 3 introduces the main types of WiFi technologies and data measures used in indoor positioning systems. Section 4 focuses on the general concept of deep learning and neural networks and discusses further the main types of neural networks that are used in the covered papers. Section 5 overviews the common evaluation metrics adopted in this review. A taxonomy of positioning system categories will be presented in Section 6 and Section 7 where each category will be investigated thoroughly and separately in these two sections. Finally, Section 8 concludes the review and outlines the future perspectives.

2. WiFi fingerprinting

WiFi fingerprinting is the most popular approach in WiFi-based indoor positioning which starts with establishing a database containing WiFi signals collected at every reference point in the targeted indoor environment. The aim of WiFi fingerprinting is to match the real-time WiFi signals received by the user with those in the database, so that positioning estimation of the user's current location could be generated based on their relevance. Since the propagation of the WiFi signal could be severely affected by the complex indoor environment, each location in the targeted area will have its own distinguishing WiFi signal pattern, or WiFi fingerprint. The more complicated the environment is, the more distinct this WiFi fingerprint will be. Thus, positioning systems could take advantage of such features of the WiFi fingerprint to accurately perform location estimation of the user.

Normally, there are two phases in the WiFi fingerprinting method, off-line phase and on-line phase. An example of the basic structure of WiFi indoor fingerprinting system is shown as Figure 1. The off-line phase is the preparation phase, in which the data is collected and preprocessed before being stored in the dataset. In an indoor positioning system, samples in the dataset are collected in a certain environment and labelled with their targets, either the buildings/floors/grids they belong to, or their exact ground

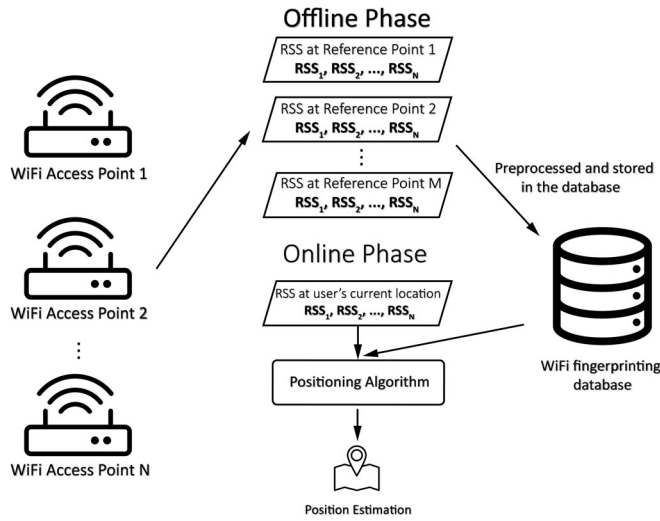


Figure 1. The basic architecture of a classic WiFi indoor fingerprinting system with machine learning as its positioning algorithm. This system has two phases: the off-line phase and the on-line phase. In the off-line phase, the WiFi fingerprinting signals, which are WiFi RSS data here, are collected, preprocessed, labelled and stored in the database. In the on-line phase, the RSS signals received by the user are compared with the signals in the database by the machine learning positioning algorithm to get the final location estimation. The basic architecture of deep learning based WiFi indoor fingerprinting system will be explained in Figures 9 and 19.

truth coordinates. To better extract the useful and meaningful information of the data, many researchers apply preprocessing methods including normalization, filling missing data, access points selection, augmentation, and calibration. Some practitioners even apply machine learning methods to extract the most powerful features of the data while reducing the dimension of the data and computational complexity of the prediction. Furthermore, in the off-line phase, the positioning prediction algorithms are pre-trained with the collected data to learn the relationship between the input data and the output prediction. In the on-line phase, a user or a receiver reports the detected WiFi signals at an unknown location to the positioning system, the system uses the same preprocessing method to filter the new data. Then, the data of the same format as in the off-line phase is fed to the positioning algorithm. Finally, the estimation of the user's current location is predicted by the positioning algorithm.

3. WiFi signal measures

This section introduces the main types of technologies and signal measures for the WiFi indoor positioning systems covered in this review.

3.1. Received signal strength (RSS)

WiFi received signal strength (RSS) technology is the most popular one used in WiFi indoor positioning. This technology uses the signal strength received by the user to

Table 1. An example fraction of the WiFi RSS data.

WAP001	WAP002	...	WAP520	LONGITUDE	LATITUDE	FLOOR	BUILDINGID
100	100	...	100	-7536.62	4864934.23	2	1
-75	-66	...	-93	-7510.44	4864949.25	3	2
100	100	...	-89	-7564.20	4864887.19	3	2
-58	100	...	100	-7336.70	4864764.48	2	1
100	-81	...	100	-7406.06	4864788.22	3	2
-77	100	...	-90	-7385.37	4864776.74	3	2
100	100	...	-92	-7391.08	4864779.91	3	2
-90	100	...	100	-7349.28	4864758.81	3	2
-74	-83	...	100	-7502.45	4864884.18	3	2
100	100	...	100	-7474.58	4864866.87	3	2

Notes: It consists of the RSS signals received from all nearby WiFi access points. The columns of FLOOR and BUILDINGID indicate the labels of the RSS data while the columns of LONGITUDE and LATITUDE represent the 2D coordinates of the location where RSS signals were collected.

estimate the user's location. Generally, WiFi RSS database contains RSS values from different access points collected at each location and the labels of the data. An example of WiFi RSS data is shown in Table 1 where values in column WAP001 to column WAP520 represent the RSS signals received from the specific WiFi access point (e.g. the values below the heading WAP001 represent RSS signals from the No.1 access point). Each row represents a reference point where its RSS signals were collected. The unit of the RSS is dBm. The value of 100 means that the RSS from the specific access point could not be received at the reference point. The columns of FLOOR and BUILDINGID indicate the labels of the RSS data while the columns of LONGITUDE and LATITUDE represent the 2D coordinates of the reference points. Trilateration, approximate perception, and fingerprinting are possible approaches to utilizing WiFi RSS for indoor positioning. Trilateration, more like GPS, makes use of three or more access points and the distance between the receiver and transmitters to calculate the possible location of the user. Approximate perception is much simpler, it estimates the final location based on the access point that gives the strongest WiFi RSS. These two methods do not require machine learning methods, so they are not included in the scope of this review.

Fingerprinting technology is arguably the most popular of the three methods and is widely used in indoor positioning systems especially those using deep learning methods. Since WiFi signals are easily attenuated and reflected in complex indoor environments, different locations could receive largely distinct RSS from multiple access points. Such a particular distribution of the RSS signals in a specific location is regarded as the fingerprint of this location. Like identifying a person by his or her fingerprint, this technology utilizes the uniqueness of the RSS at a specific location as the estimation evidence to predict where the user is. Thus, the key task of the fingerprinting method is to match the RSS detected at a location with the RSS collected in the database.

3.2. Channel state information (CSI)

Channel state information (CSI) is another rich information that could be derived from WiFi signals through orthogonal frequency-division multiplexing (OFDM). CSI is the representation and description of the WiFi channel properties in a communication link

between the access points and the receiver. This representation reveals the combined effect of multipath, scattering, fading, and power decay with distance during the propagation of the WiFi signal (Basri & El Khadimi, 2016; He & Chan, 2015). Due to its nature, the CSI is more stable than the RSS on a timescale but has strong specificity over space. In addition, a single antenna of the WiFi transmitter has many subcarriers and the features of the subcarriers are different in different antennas. So that is the reason why researchers aim to use CSI to achieve sub-metre level accuracy in indoor positioning systems. CSI signals could be divided further into two kinds: amplitude and phase. Both of them could be used as input for the indoor positioning system. However, the CSI data is rather harder to get than WiFi RSS. Unlike RSS that could be easily obtained from the receiver, the CSI data need to be derived from the driver of the WiFi receiver on a laptop. Therefore, implementing a CSI-based indoor positioning system on smartphones is more of a challenge.

3.3. Round-trip time (RTT)

WiFi round-trip time (RTT) information is the creation of the fine time measurement (FTM) protocol for ranging proposed by the IEEE 802.11-2016. It is a new protocol that could be used to directly calculate the time duration of a single WiFi signal to travel from the transmitter to the receiver. However, to the best of our knowledge, there is only one research paper in this area that uses both RTT and deep learning in the indoor positioning system.

4. Deep learning neural networks

Since all the papers covered in this survey are based on deep learning and neural networks, it is essential to first understand them both. This section introduces the concept of deep learning and the several main types of neural networks that are used in the scope of this survey.

4.1. Deep learning

Deep learning is considered as an evolution of machine learning. It is based on neural networks but more focused on deeper representation learning. Deep learning neural networks learn the increasingly meaningful representations from data via multiple layers to make predictions (Chollet, 2018). Layer is a processing stage or unit that uses a specific function to extract information from the input to this layer and then outputs the higher-level information to the next layer. Because a layer is the basic computing unit of deep learning, the number of layers could be used to describe the 'depth' of deep learning.

The models used to learn these representations are neural networks. A simple neural network consists of three types of layers: input layer, hidden layer, and output layer. Figure 2 shows a basic structure of a neural network. The input layer contains the input data of the neural network. The output layer is the exact layer that generates the output from the representations learned via previous layers. Hidden layers are the main computing part of the neural network where the meaningful higher-level representations of the input data are learned. The most obvious difference between deep learning and

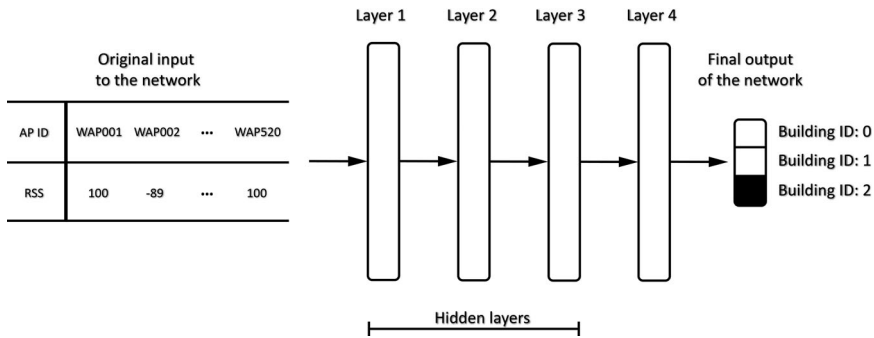


Figure 2. The structure of a deep neural network based on WiFi RSS for predicting the building where the user is in. In this structure, the input layer contains the original input. Layer 1 to layer 3 are the hidden layers. Layer 4, between hidden layers and the final output, is the output layer.

simple neural networks is that the networks in deep learning usually have more layers and more complicated structures than simple neural networks. Recently, advanced neural networks can have tens or even hundreds of hidden layers.

In this review, the performance of more than 150 WiFi indoor positioning systems using neural networks, including deep neural networks and simple neural networks, will be compared. The effect of using different neural networks and their complexity on WiFi indoor positioning will be investigated. Thus, different types of neural networks will be covered briefly in the following subsections, while the number of hidden layers (i.e. the same concept of 'depth' for a neural network) will be used to compare the complexity of different deep learning methods in Sections 6 and 7.

4.2. Artificial neural network (ANN)

Artificial neural network (ANN) is a general and basic type of neural networks, see Figure 3. An ANN is based on a collection of connected units or nodes called neuron. The model of a single ANN neuron is shown on the right of Figure 3 where it stores the input data or information learned from previous layers and passes it through to the next layer. The

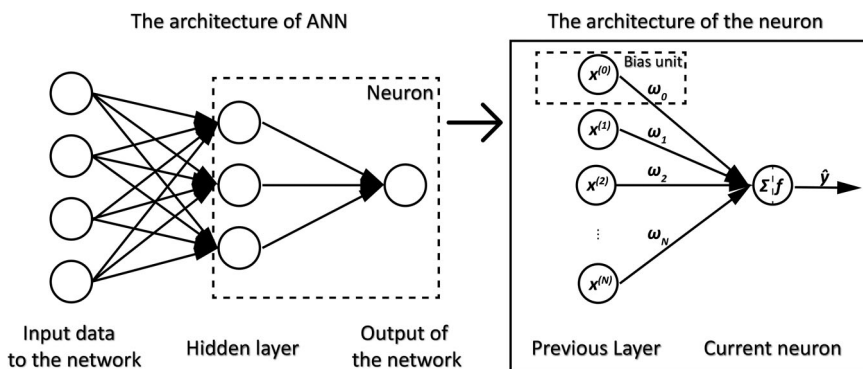


Figure 3. The basic architecture of ANN and its neuron.

output $\hat{y}$ of the neuron is defined as

$$\hat{y} = f \left(\sum_{i=1}^N \omega_i x^{(i)} + \omega_0 \right) \quad (1)$$

where N represents the maximum number of the neurons in the previous layer, $\hat{y}$ is the output of this neuron, $x^{(i)}$ stands for the information stored in the i th neuron of the previous layer, $x^{(0)}$ is the bias unit that is set to 1, ω_i and ω_0 are the weights learned by the neural network where ω_0 is the weight of the bias unit, and f is the activation function that generates the output based on the input information and the weights from all connected neurons in the previous layer. In the output layer, such functions are responsible for performing different prediction tasks like regression or classification.

ANN was first used to describe the simply structured network that only has an input layer, one hidden layer and an output layer. Then several changes were made to ANN and named differently. They are multi-layer perceptron (MLP), deep neural network (DNN), back propagation neural network (BPNN), feed forward neural network (FFNN), extreme learning machine (ELM), parallel multilayer neural network (PMNN), etc. To be clear and specific when making comparisons in the following sections, all these neural networks that are similar to ANN will be included in the group of ANN. WiFi indoor positioning systems generally employ ANN directly on the preprocessed input data. Due to its simplicity, ANN only aims to find the mapping from the numerical WiFi data to the specific location.

4.3. Auto-encoder (AE)

Auto-encoder (AE) is an unsupervised learning neural network. The common structure of an AE is shown in Figure 4. This network has mainly two parts, the encoder part and the decoder part. The encoder part takes the input data into a neural network and uses an unsupervised method to learn the compact representation of the data. The decoder part decodes such compact representation so that the output is as similar to the original input as possible.

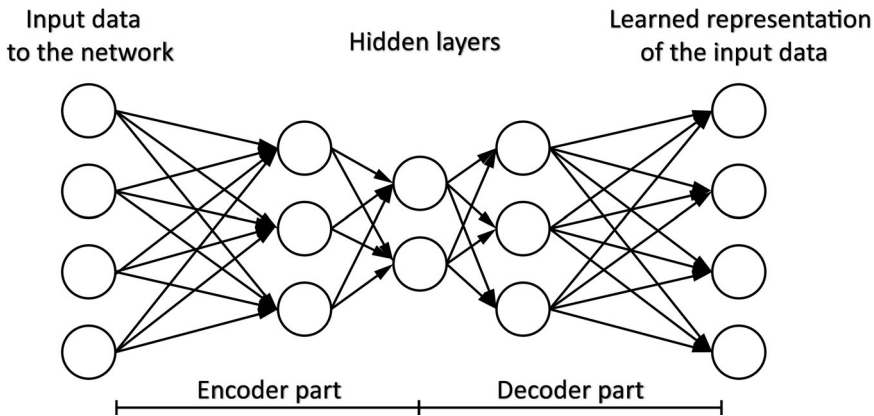


Figure 4. The structure of an AE network where the encoder part compresses the input data, while the compressed data is decoded by the decoder part.

AE also has many variations such as denoising auto-encoder (DAE), stacked auto-encoder (SAE), and stacked denoising auto-encoder (SDAE). All these variations are included in the group of AE in the following sections. Like implementing traditional unsupervised machine learning method, indoor positioning systems utilizing AE are expecting a filtered and refined version of the input WiFi data and to see if there are hidden connections between the compressed input data and the positioning estimation. By doing so, both the complexity of such high-dimensional WiFi data and irrelevant information of the sparse data could be reduced.

4.4. Convolutional neural network (CNN)

Convolutional neural network (CNN) is a neural network famous for its ability to make image classifications. As shown in Figure 5, the main features of this network are that the input is mainly two-dimensional image data, and the layers of CNN use the convolution operations to summarize the presence of features in an input image. Such layers using convolution operations are called 'convolutional layers' which extract the higher-level information from its input image data. The convolutional layer utilizes a small filter that slides over the input at all possible locations while getting a specific value at every location. Then these values from all possible locations are transformed into a new 'image' of data which is then fed to the following layer.

Furthermore, CNN contains special layers like max pooling layer and flatten layer to better extract features from the input image data (see Figure 5). Note that the CNN dealing with 1D data is called 1D-CNN. Hierarchical features of the input are the main purpose of using CNN in WiFi indoor positioning. To imitate the way people detect certain semantic patterns in images, indoor positioning systems are trying to seek for such patterns in WiFi data with the help of CNN. Converting WiFi signals to 2D images, simply forming 2D vectors of WiFi data or implementing 1D CNN on WiFi data are the three most common ways of employing CNN in WiFi indoor positioning.

4.5. Recurrent neural network (RNN)

Recurrent neural network (RNN) utilizes the layer called 'recurrent layer' to process the sequence data. As a result, RNN is more likely used to perform tracking in WiFi indoor

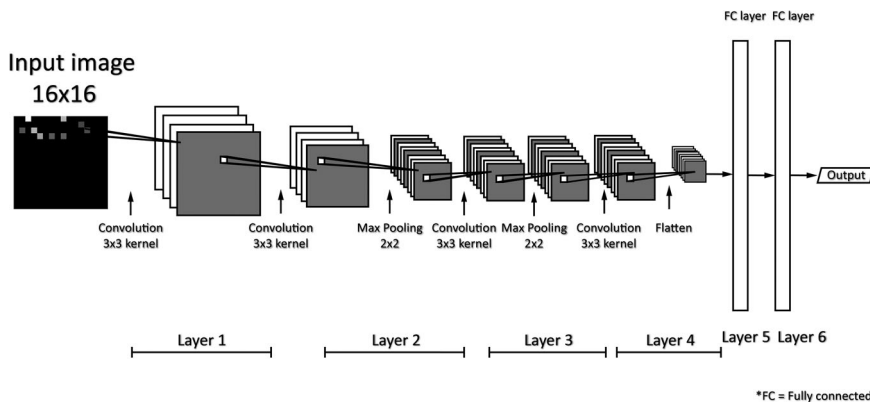


Figure 5. A real example of CNN architecture introduced in Sinha and Hwang (2019).

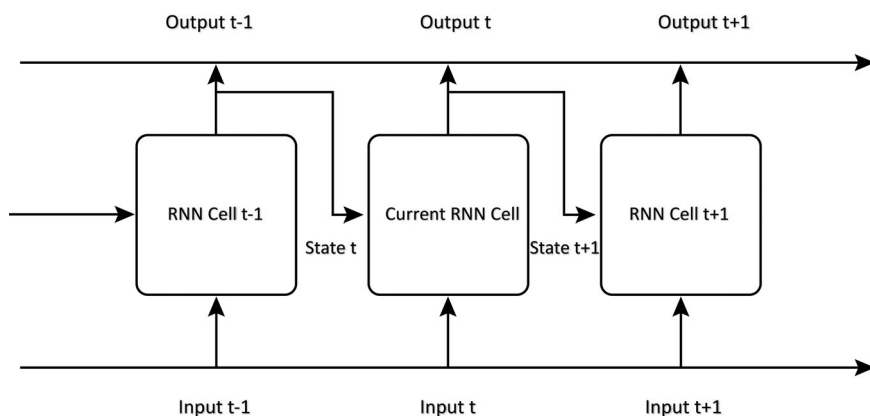


Figure 6. The basic structure of RNN. The recurrent layer utilizes the current input element *input t*, and the state from the last layer *state t*, and then it generates a temporary output *output t* and a new state, *state t+1*, which represents the information of what it has seen so far. Through the recurrent layers, RNN is able to extract features from time-series data or sequence data.

positioning scenarios. The basic structure of RNN is presented in Figure 6. The way RNN learns the representations of the data is like how we read a sentence. It iterates all sequential elements in the data and learns the hidden correlations among them. The RNN cell in the recurrent layer takes advantage of the current input element *input t* and the state from the last RNN cell *state t* and then generates a temporary output *output t* and a new state *state t+1*. The *state t+1* represents the information of what the RNN has seen so far.

However, the basic structure of RNN has many drawbacks such as gradient vanishing and exploding problems, which is to say that RNN is easy to forget the information at the beginning of the sequence data. Long-short term memory (LSTM), an extended version of RNN, is then introduced to solve this problem using the structures of forget gate, input gate and output gate. In the following sections, both RNN and LSTM will be included in the group of RNN. Due to RNN's advantage of analysing time series data, systems utilizing such networks are focusing on collecting continuous WiFi data with time step in a certain period of time. Based on assessing the user's motion in the time period, systems could better estimate the location than other networks under the same circumstances. Therefore, movement tracking or predicting the user's walking trajectory are the main tasks of RNN in WiFi indoor positioning.

4.6. Other neural networks

There are several neural networks that are included but not widely used in the scope of this survey. They are deep belief network (DBN), generative adversarial network (GAN) and capsule neural network. DBN is formed by multiple stacked restricted Boltzmann machines (RBMs) and is using a greedy learning strategy to generate the probabilistic distribution among the input and hidden layers (Hinton, 2009; Kozma et al., 2018). GAN, as shown in Figure 7, is the type of neural network that learns to simulate data. It consists of two models: a generative model and a discriminative model. These two models are

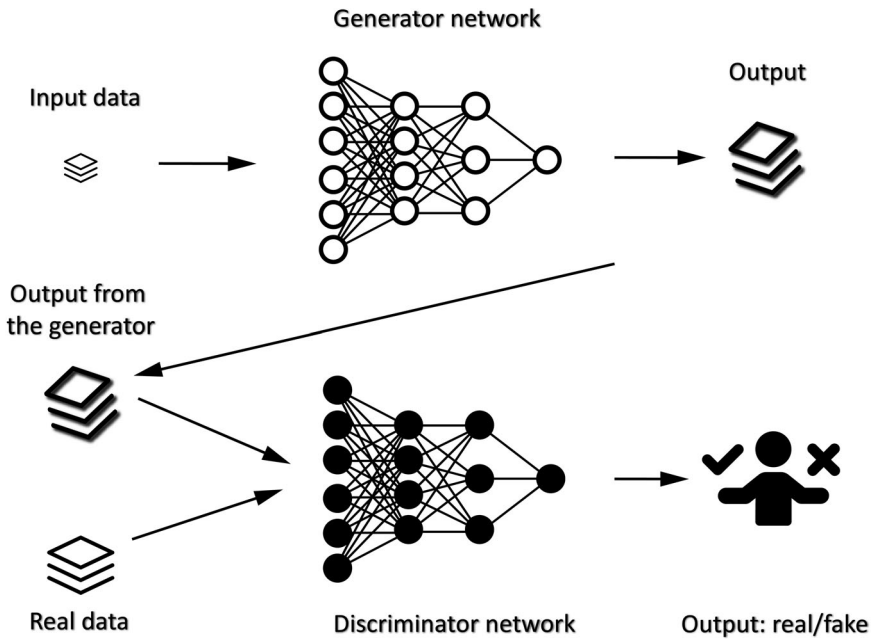


Figure 7. The basic structure of GAN. The generator and the discriminator are trained at the same time. The generator network generates data as similar to the real data as possible and the discriminator network outputs the similarity between the real data and the generated data.

trained simultaneously while the generative one aims to generate data as similar to the real data as possible and the discriminative one outputs the similarity between the real data and the generated data. Capsule neural network could be regarded as a simpler version of CNN but only needs fewer computational costs.

5. Evaluation metric

The aim of the WiFi indoor positioning system is to accurately locate the user. Ideally, a positioning system predicts the user's location in a 3D space giving the result of 3D coordinates. However, because of the challenge in the signal similarity across different floors, most people only consider the positioning in a 2D space. Even in the 2D space, some estimate the exact 2D coordinates of the location while others who divide the testbed into several grids only predict which grid the user is on. To cope with this situation, researchers will offer floor estimation and even building estimation at the same time. Combining the building and floor predictions with the 2D positioning estimation, the user's accurate location in the 3D space could thus be deduced.

Among all the papers reviewed, there is no general set of evaluation metrics. This lack of a convincing general evaluation method is caused by several reasons. Firstly, though UJIIndoorLoc by Torres-Sospedra et al. (2014) is a commonly acknowledged public WiFi dataset in indoor positioning field, it only focuses on WiFi RSS signals. Many researchers developed their systems based on the CSI signals in order to achieve sub-metre level accuracy. As a result, there is no such public WiFi CSI dataset for indoor positioning

which leads to a diversity of testbeds and datasets among all the CSI papers. Even systems based on WiFi RSS signals may not use the public data as the criterion to evaluate their performances. Secondly, practitioners in this area do not follow a general theme, they try to consider the positioning task differently which results in needs for distinct evaluation frameworks. In WiFi indoor positioning, according to different prediction aims, there are mainly two types of systems: one regards the indoor positioning problem as a classification problem, the other treats the problem as a regression one. Thus, for this oil and vinegar situation, several evaluation metrics are considered.

For the benefit of getting a comprehensive result, a standard set of metrics is used to evaluate and compare the performance of most deep learning based systems. This section will introduce six general evaluation frameworks for WiFi indoor positioning systems which are commonly used among the reviewed papers. These metrics are hitting rate, Mean Distance Error (MDE), Root Mean Squared Error (RMSE), Cumulative Distribution Function (CDF), Complexity (i.e. the number of hidden layers) and Testing Time. Sections 6 and 7 will include these metrics for the evaluation and comparisons of all covered systems. Specifically, hitting rate is the main evaluation criterion for classification systems, MDE and RMSE are quantitative criteria for regression systems. CDF is another evaluation method for regression systems. But due to the nature of CDF, systems using this as the performance evaluation method could not provide direct, valid and convincing results for comparisons. The evaluation metrics of complexity and testing time offer another perspective to assess the feasibility of the indoor positioning systems.

5.1. Hitting rate

For the WiFi indoor positioning classification systems, research papers could be divided into two major groups. The first group aims to achieve building-level accuracy or floor-level accuracy, thus their classification goals are to locate the objects or users to a specific building or floor. To enhance the accuracy, the other group further refines their classification output classes to smaller zones or grids in the testbeds. In this way, the indoor positioning problem turns to predicting which grid or zone the targeted object is in. But the main challenge here is that due to the difference in setting the size of the grids or zones, it could be hard to fairly compare the performances of all distinct classification systems. The major evaluation metric among all the covered papers is the hitting rate which represents the prediction accuracy in the classification problems.

The hitting rate is defined as

$$\text{Hitting rate} = \frac{\text{the number of corret predictions}}{\text{the number of all predictions}} \times 100\% \quad (2)$$

With such an evaluation method, a general expression of the performance of classification systems could be derived. To be fair, all floor-level classification systems will be compared first and then all zone/grid-targeted systems will be compared.

5.2. Mean distance error (MDE)

The most direct way to qualify a regression indoor positioning system is to judge it by its mean error. Most WiFi indoor positioning regression systems use a regression layer

as the output layer to predict the exact coordinates of the targeted users or objects. This prediction is commonly based on the prior assumption or confirmation of which floor the user or object is on. Therefore, the regression systems either test their performances on a single-floor testbed or use a floor-level classifier first and then form a regressor based on the classification results. There are 2 main evaluation metrics for regression systems, they are Mean Distance Error (MDE) and Root Mean Squared Error (RMSE). In Sections 6 and 7, the regression indoor positioning systems based on MDE and RMSE will be compared separately due to the evaluation method they used.

MDE is obtained from the mean value of all distance errors. The common distance error is the Euclidean distance between the predicted coordinates and the ground truth coordinates of a specific location. The MDE is defined as

$$MDE = \frac{1}{N} \sum_{i=1}^N Dist_i, \quad (3)$$

$$Dist_i = \sqrt{[(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2]}, \quad (4)$$

where N is the total number of test samples, $Dist_i$ is the Euclidean distance between the predicted coordinates $(\hat{x}_i, \hat{y}_i)$ and the ground truth coordinates of the i th sample (x_i, y_i) .

5.3. Root mean squared error (RMSE)

RMSE is another metric that is commonly employed to assess the performance of a regression system. RMSE is the standard deviation of the errors between the predictions and the true values, which is defined as

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2]} \quad (5)$$

where N is the total number of test samples, $(\hat{x}_i, \hat{y}_i)$ are the predicted coordinates of the i th sample and (x_i, y_i) are the ground truth coordinates of the i th sample.

5.4. Cumulative distribution function

Cumulative distribution function $F_X(x)$ shows the probability that X will have a value less than or equal to x . Some regression systems use CDF other than MDE or RMSE as the evaluation metric to describe their performances. For instance, the result of a system using CDF would be that it could achieve a distance error of 2 m with the probability of 90%. Note that the ways researchers use CDF to evaluate their systems are different, i.e. they present their distance errors with distinguishing probabilities. So comparing systems using such evaluation metric will not be the main focus of this survey. Readers could find records of CDF in Tables 2 and 4.

Table 2. Comparison of the covered WiFi-based indoor positioning systems using Deep Learning as a feature extraction method.

Research work	Input	Deep learning method	No. of hidden layers	MSE (m)	RMSE (m)	Hitting rate (%)	Data size	Test time (ms)
M. Zhang et al. (2021)	RSS PDR (pedestrian dead-reckoning)	LSTM	20	0.42	–	–	24,280	309
JunLin et al. (2020)	RSS	SAE	7	–	–	Floor: 95.139	19,937	–
C. Liu et al. (2020)	RSS	DNN+SDAE	2	2, 80% chance	–	–	–	–
X. Wang et al. (2020)	RSS	DGP	1	1.668	–	–	–	–
Y. Wang et al. (2020)	RSS	DNN + SDAE	3	UJIIndoorloc: 5.29; real-dynamic: 1.72 real-static: 1.22	–	Floor: 94.60	UJIIndoorloc: 19,937/ real scenario: 19,800	–
Xue et al. (2020)	RSS	BPNN	1	0.87	–	–	–	–
L. Zhang et al. (2020)	RSS	DNN	4	–	1.79	–	428,000	–
Abbas et al. (2019)	RSS	SDAE	4	larger testbed: 2.64; small testbed 1.21	–	–	building 4320; office 1200	–
P. Dai et al. (2019)	RSS	DNN	2	1.39	–	–	1000	–
Hoang et al. (2019)	RSS	RNN+LSTM	2	0.75	–	–	365,000	–
Hsu et al. (2019)	RSS	DBN	7	265% chance	–	–	4500	–
Ma et al. (2019)	RSS ranking	CNN	4	1.54	–	–	2260	–
Own et al. (2019)	RSS	CNN Capsule Network	4	0.99	–	–	–	850
Qian et al. (2019)	RSS	CNN+RNN +MDN	5	–	6.25	–	–	–
Y. Shao et al. (2019)	RSS	ANN	1	–	–	Floor: 93.84	697	–
Soro and Lee (2019)	RSS	DNN	4	0.68/0	–	–	2100/8460	–
P. Wu et al. (2019)	RSS	MLP	1	less than 2	–	–	1071	–
G. Zhang et al. (2019)	RSS	CNN	7	0.9554	–	–	–	–
Dou et al. (2018)	RSS	deep Q	2	0.55, 75% chance	–	–	19,937	–
Elbakly et al. (2018)	RSS	MLP	1	–	–	Mall Floor: 91.8 UJIIndoorloc: 85	–	–
Kim, Lee, et al. (2018)	RSS	DNN+SAE	5	9.29	–	building: 99.82 ; Floor: 91.27	13,981	–
Le et al. (2018)	RSS	DBN	2	–	6.2	–	5249	–
Research work	Input	Deep learning method	No. of hidden layers	MSE(m)	RMSE(m)	Hitting rate (%)	Data size	Test time (ms)
J. Liu et al. (2018)	RSS	MLP+AE	5	1.29–1.30	–	–	9000	–

D. V. Nguyen et al. (2018)	RSS	EBNN	–	0.62	–	–	–	–
Qi et al. (2018)	RSS	ensemble ELM	10	–	–	96	350	–
Soro and Lee (2018)	RSS	MNN	15	–	office: 1.39; corridor: 1.50; UJIIndoorloc: 8.63	–	N/A, N/A, UJIIndoorloc: 19,937	–
Zou et al. (2018)	RSS	DNN+CNN+AE	13	–	1.48	–	–	–
Dinh-Van et al. (2017)	RSS	EBNN	–	2.25	2.8	–	–	–
X. Wang (2019)	CSI amplitude & phase	CNN	34	0.98	–	–	15,000	587
C. Zhou and Gu (2017)	RSS	AE+DNN	–	–	1.73	–	–	–
D. V. Nguyen et al. (2016)	RSS	DNN	–	5.77	6.85	–	–	–
W. Zhang et al. (2016)	RSS	DNN + SDAE	3	–	0.34	–	7280	250
Bernas and Placzek (2015)	RSS	ANN+FCNN	1	1.08	–	–	6300	70
Campos et al. (2014)	RSS	ANN	2	1.7	–	97	175,560	–
W. Zhang et al. (2014)	RSS	DNN	4	–	–	98.5	49,140	–
Y. Xu and Sun (2012)	RSS	MLP	1	1.26	–	–	54,600	–
Laoudias et al. (2009)	RSS	RBFFNN	1	3.4	–	–	3210	–
B. F. Wu et al. (2007)	RSS	ANFIS	1	2	–	–	17,000	–
Elbakly and Youssef (2020)	RSS image	CNN(VGG19)	24	–	–	Floor: 92	28,150	–
Haider et al. (2019)	RSS image	CNN	4	4, 95.94% chance	–	93.701	–	2.03
W. Shao, Luo, Zhao, Wang, et al. (2018)	RSS+magnetic field	CNN+MLP	18	less than 1 m	–	–	–	–
M. Liu et al. (2017)	RSS + magnetic field + heading	CNN	10	1.32, 95% chance	–	scene: 89.8	–	–
Belmonte-Hernández et al. (2019)	RSS of WiFi+BLE +Xbee+IMU +magnetic	ANN	2	0.45	–	93	–	3.689
Y. Li et al. (2019)	RSS + IMU	MLP	2	2.7	3.3	–	12867	0.75
Research work	Input	Deep learning method	No. of hidden layers	MSE (m)	RMSE (m)	Hitting rate (%)	Data size	Test time (ms)

(Continued)

Table 2. Continued.

Research work	Input	Deep learning method	No. of hidden layers	MSE (m)	RMSE (m)	Hitting rate (%)	Data size	Test time (ms)
Z. Zhou et al. (2020)	CSI	CNN	106	1.5	–	–	–	PC: 210; Phone: 1500
G. S. Wu and Tseng (2018)	CSI	DNN+SAE	4	1, 50% chance	–	–	10,000	–
T. Li et al. (2018)	CSI amplitude	DNN	3	0.6-0.7	–	–	58,800	–
X. Wang, Gao, Mao, and Pandey (2016)	CSI amplitude	SAE	4	living room: 0.95; lab: 1.8	–	–	living: 19,000; lab: 50,000	–
X. Wang, Gao, Mao, and Pandey (2015)	CSI amplitude	SAE	4	living room: 0.95; lab: 1.8	–	–	living: 19,000 lab 50,000	–
H. Chen et al. (2017)	CSI amplitude image	CNN	4	1.37	–	91.78	537,600	–
X. Wang, Gao, and Mao (2017)	CSI amplitude & AOA	bi-modal AE	3	lab: 1.57; corridor: 2.15	–	–	15,000	lab: 665.3; corridor: 544.0
X. Wang, Gao, and Mao (2016)	CSI phase	SAE	3	living room: 1.08; lab: 2.01	–	–	living: 15,200; lab: 40,000	living room: 378.0; lab: 377.0
X. Wang, Gao, and Mao (2015)	CSI phase	SAE	3	living: 1.08; lab: 2.01	–	–	living: 15,200 lab: 40,000	380
X. Wang et al. (2018b)	CSI amplitude & phase	AE+Residual Neural Network	0.86	–	–	15,000	–	–
X. Wang, Wang, and Mao (2017b)	CSI amplitude & phase	CNN	34	0.98	–	–	15,000	587
J. Li et al. (2016)	CSI amplitude & phase	BPNN	1	1.57	–	–	500,000	–
H. Li et al. (2019)	CSI image	CNN	7	1.58	–	–	–	–
Zhao et al. (2019)	CSI image	CNN	5	NLoS: 1.11; LoS: 0.46	–	–	2,268,000	–
X. Wang, Wang, and Mao (2017a)	CSI image	CNN	8	lab: 1, 40% chance; corridor: 3, 60% chance	–	–	lab: 14,400; corridor: 9600	–
X. Wang et al. (2018a)	CSI AOA image	CNN	8	lab: 1.79; corridor: 2.29	–	–	lab: 1500; corridor: 10,000	–

5.5. Complexity (i.e. the number of hidden layers)

To better investigate the effectiveness of the covered systems, the complexity of the neural networks will also be analysed. Though the complexity of a neural network includes the framework of the neural network model, the size of the model, optimization process and data complexity (Hu et al., 2021), few covered papers provide such detailed information which makes it hard to conduct a deeper analysis and comparison from this perspective. The number of hidden layers will be used as the metric to represent the complexity of the deep learning approaches employed by WiFi indoor positioning systems. Hidden layers are the layers between input layer and output layer in deep learning neural networks as shown in Figure 8. Since all these more than 150 covered papers are making use of deep learning methods, it is significant to bring out complexity as another dimension to compare all the systems with. And as users are relying more on their smartphones, it is within the near future that smartphone will be the top one device people use for indoor positioning. Thus, taking the complexity of the systems into consideration is urgent and necessary. Furthermore, the number of hidden layers is general information that most papers would offer to demonstrate the complexity and computational cost of their systems. Therefore, such a metric is introduced to make the comparison among WiFi indoor positioning systems using deep learning methods.

5.6. Testing time

Another important goal of all the indoor positioning systems is to locate the object or user quickly. Testing time is the amount of time required by the indoor positioning system to perform all necessary computation/prediction on a single testing sample. Either for tracking or navigation, it is essential to perform the indoor positioning instantly as the object or user is probably moving at a low but not neglected speed. Being slower means that the prediction is getting further away from the user's exact current location. From the computational cost perspective, the testing time is also a conclusive metric to evaluate a certain system. While most papers use laptops as the receiver of their indoor positioning systems, some are implementing the positioning algorithms on the smartphones. This

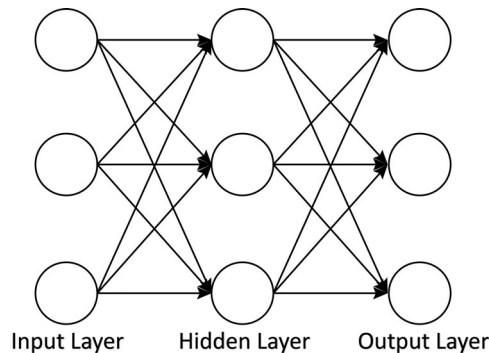


Figure 8. Hidden layers are the layers between input layer and output layer. The number of hidden layers could vary from only one to hundreds.

trend indicates that researchers are focusing on using such portable device to perform the positioning. Thus, introducing testing time as an evaluation metric would help readers understand the potential of a system in the near future. All the testing time results from the covered papers are included to offer a new perspective to evaluate the feasibility of the indoor positioning systems.

6. Deep learning-based feature extraction methods for WiFi-based indoor positioning

This section overviews those WiFi indoor positioning systems that use deep learning merely as a feature extraction method. This means that deep learning methods are only used to find more effective representations of the input data. These systems may use probabilistic methods or traditional machine learning methods as the final prediction algorithms. Thus, to compare these indoor positioning systems, their positioning performances, computational complexity and costs are to be investigated. In other words, the number of hidden layers and testing time are included evaluation metrics other than hitting rate, MDE or RMSE.

The basic structure of the systems reviewed in this section are illustrated in Figure 9. Input data, including WiFi RSS, CSI and hybrid signals from other sensors, are preprocessed before being fed into the feature extraction methods. For the best performance of the deep learning feature extraction methods, the input data are to be normalized, calibrated, augmented, classified or preprocessed with dimensionality reduction using statistical methods or traditional machine learning algorithms like SVM and PCA during preprocessing. Deep learning neural networks are used to extract hierarchical features of the input data. And the types of neural networks included in this section could be classified as Artificial Neural Networks (ANN), Convolution Neural Network (CNN), Auto Encoder (AE), Deep Belief Network (DBN), Recurrent Neural Network (RNN) and other architectures (e.g. Deep Gaussian Process (DGP)).

To give a comprehensive view of the trends in deep learning-based feature extraction methods for WiFi indoor positioning systems, networks that are used less than 5 times will not be classified as a separate type in the comparison part. And networks that are simple variation of the ANN will be classified as ANN for the convenience of comparison. They include Multilayer Perceptron (MLP), Deep Neural Network (DNN), Back Propagation Neural Network (BPNN), Extreme Learning Machine (ELM), Feed-forward Neural Network (FFNN), Discriminant-Adaptive Neural Network (DANN). Also, Stacked Autoencoder (SAE) and Stacked Denoising Autoencoder (SDAE) are classified as AE while Long Short-Term Memory (LSTM) is classified as RNN.

Table 2 compares the WiFi indoor positioning systems that use deep learning as feature extraction methods. The evaluation metrics are the number of hidden layers, MDE, RMSE, hitting rate, data size and testing time. All the information about the systems is derived from the papers directly and the missing entries imply that the papers do not have the specific information. The positioning performance is clearly demonstrated by MDE, RMSE, hitting rate or CDF, while the feasibility and computational complexity could be deduced by the number of hidden layers, data size and testing time. A more detailed and refined comparison can be found in Section 6.5.

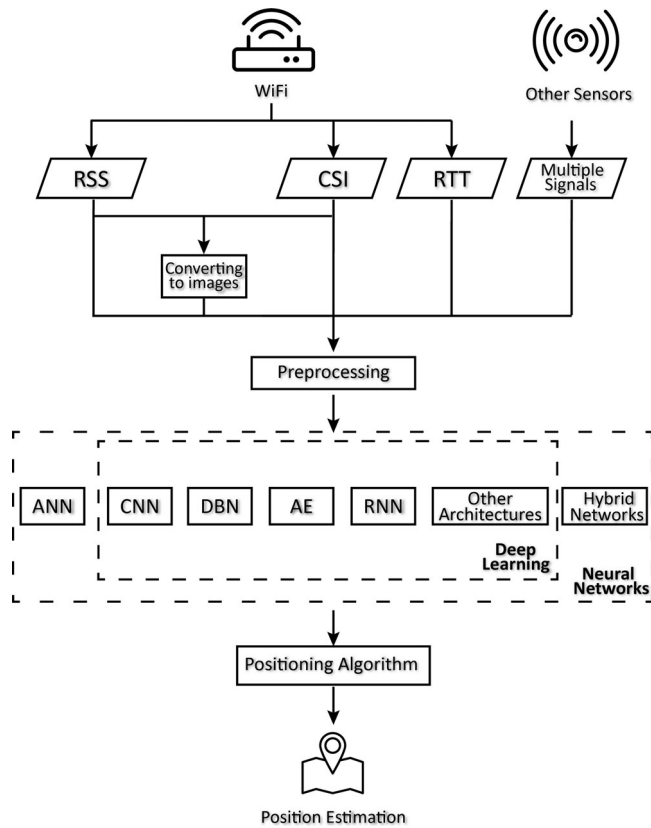


Figure 9. The general process of WiFi-based indoor positioning systems employing deep learning as a feature extraction method. Please note that the difference between deep learning and neural networks is identified in this figure. While in the comparisons, all the different neural networks adopted by covered systems will still be compared whether they are deep neural networks or simple neural networks.

6.1. ANN

Qi et al. (2018) proposed a system that adopts several ELM (a variation of ANN) classifiers' outputs to estimate the user's location as shown in Figure 10. In the off-line phase, the system first utilizes principal component analysis (PCA) to perform dimension reduction on the RSS data to improve performance and filter out irrelevant information from the training data. The preprocessed data are then fed into the ensemble model. In the ensemble model, different ELM classifiers are trained for each floor to get individual floor-level classification results. After these classifications are performed, all the results from the ELMs are passed to the final classification algorithm which is majority voting. The final floor estimation results are derived from the voting strategy. In the on-line phase, real-time RSS data are filtered through PCA and passed to the ensemble model. The floor-level prediction is done by the final classification function. The testbed of this system is the student dormitory of Nanjing University of Posts and Telecommunications that has 7 floors and 95 APs. To verify the performance, they test the system on 700 test RSS measurement. The final floor-level prediction accuracy is more than 96%.

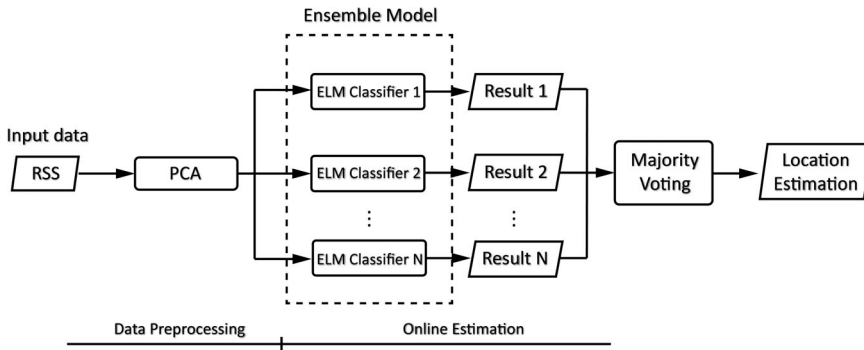


Figure 10. The structure of the system proposed by Qi et al. (2018). The RSS data are preprocessed by PCA and then fed into multiple ELM classifiers. The results from these classifiers are used to predict the location of the user.

Belmonte-Hernández et al. (2019) developed a WiFi indoor positioning system called SWiBluX based on the feature extraction method of ANN. The system of SWiBluX is a user tracking system. The novel parts of this system are its multi-source input data, the implementation of so-called multi-phase statistical fingerprint and deep learning disruptive approach and the use of Gaussian outlier filter for error reduction of the final estimation. As shown in the Figure 11, the input of the SWiBluX consists of the RSS of XBee, Bluetooth and WiFi. To deal with the instability of the RSS signals, the system adopts Step and Velocity Estimator and Yaw (i.e. the direction user is heading to) estimation. With the benefit of utilizing the information of user's movement, this system is able to better track the user and prevent from the impossible location estimation that could be deduced by merely predicting user's location based on the RSS signals. The multi-source data and the yaw information are transformed and stored in the feature vector. The ANN is then used to perform fingerprinting estimation and output the probabilities of user's being at a specific location. The Gaussian filter detects the outliers from the ANN's output and then passes the processed data to the particle filter. The particle filter estimates the user's final location based on the realistic movement model. The testbed of this system is split into 70 cells ranging from 150 cm × 150 cm to 240 cm × 240 cm where 7500 feature vectors are collected at each cell. The MSE of this system is 0.4541 m and the improvement of the positioning results is up to 50% compared to related indoor positioning systems. It is also proved by Belmonte-Hernández et al. that yaw/heading information, the Gaussian outlier filter and the particle filter could all bring huge improvement to the system.

6.2. CNN

Elbakly and Youssef (2020) proposed a floor prediction system based on RSS called 'Story-Teller'. The idea is to transform the RSS signals to the form of image and use CNN to perform basic predictions. One important theme of the system is that the system assumes the APs giving the strongest RSS are located on the floors near to the floor where the user is on. Thus, it narrows the candidate floors and only focuses on the targeted floor and the ones above or below it. The targeted floor, the floor above and the

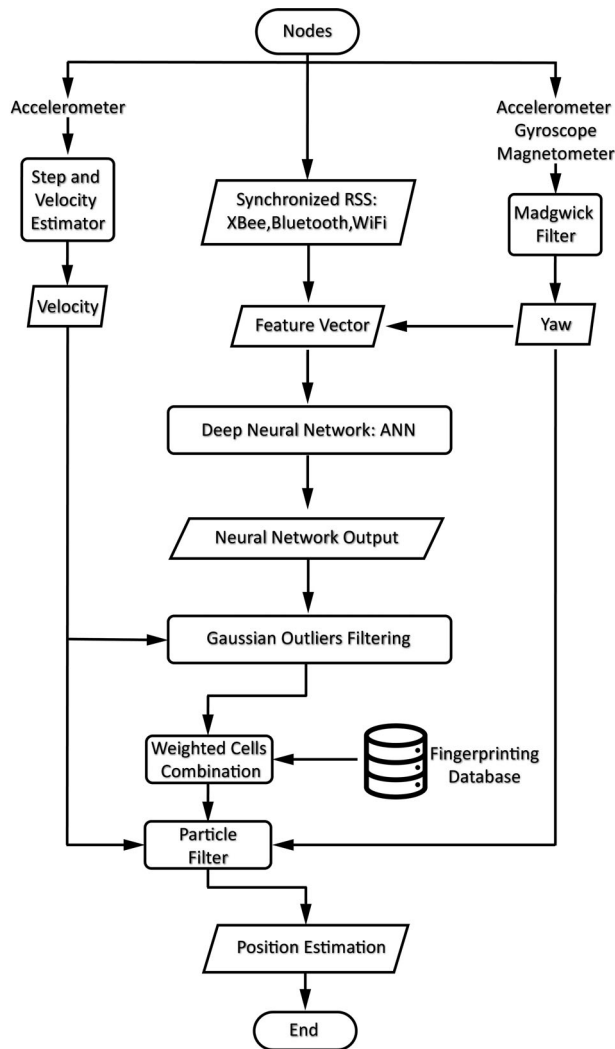


Figure 11. The structure of the system proposed by Belmonte-Hernández et al. (2019). The input data of the system are the RSS signals of XBee, Bluetooth and WiFi. The system also adopts information from Step and Velocity Estimator and Yaw estimation in the feature vectors. The ANN is then used to perform fingerprinting estimation and generate the probabilities of the user's being at a specific location. The Gaussian filter detects the outliers from the ANN's output and then passes the processed data to the particle filter. The particle filter estimates the user's final location based on the realistic movement model.

floor below together are considered as a virtual building. The prediction is implemented on this virtual building. Horizontally, the system also minimizes the area where the user could be based on the same hypothesis. The structure of the neural networks in this system is shown in Figure 12. The images of RSS signals are normalized and passed to the CNN. The CNN aims to generate a normalized floor estimation. Such estimation can then be denormalized to the range of candidate floors and be processed through weighted centroid method to get the final floor prediction.

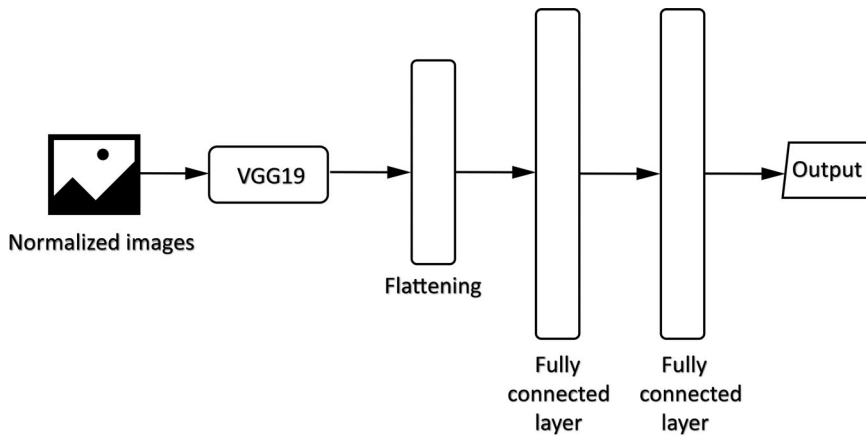


Figure 12. The architecture of the deep learning network in StoryTeller.

The system proposed by Zhao et al. (2019) also transforms the input data into images. To cope with the time-varying measurement noise and the process noise in indoor positioning problems, this system utilizes convolutional neural network and dual factor enhanced variational Bayes adaptive Kalman filter (dual factor EVBAKF). The proposed dual factor EVBAKF makes adaptive estimation of measurement noise covariance matrix (MNCM) and process noise covariance matrix (PNCM). With the help of these two matrices, the system greatly reduces the error in WiFi indoor positioning. The system first derives the CSI information from OFDM which contains the spatial and temporal features hidden in WiFi signals. The preprocessing model then compresses such CSI information to the form of image which leads to the success implementation of CNN. The testing area is of 50 m² and is split into 42 reference points and 4 target positions where 54,000 CSI measurements are collected for each reference point. The researchers also measure 3600 CSI data groups for each target position. The images of the CSI data come from the 3 receiving antennas and one transmitting antenna. CSI space-time matrix of these $3 \times 1 = 3$ antenna links are formed with the number of data packets (i.e. time factor). Every image represents 90 packets of CSI data. Thus, 600 and 40 CSI fingerprint images are obtained for each reference point and each target position, respectively. The test results indicate that the proposed system has an improvement of 22% in Line-of-Sight (LoS) environment and 9.8% in None-Line-of-Sight (NLoS) environment, compared to traditional CNN models like ResNet18, ResNet34, ResNet50 and ResNet10. And the mean squared error of the system is 1.11 m in NLoS scenarios and 0.46 m in LoS scenarios.

6.3. AE

X. Wang, Gao, Mao, and Pandey (2015, 2016) and X. Wang, Gao, and Mao (2015, 2016, 2017) all proposed similar systems with different inputs and minor details. These systems are widely cited and commonly used as reference systems to compare with the newly proposed indoor positioning systems in this field. Their main idea is using Auto-encoder to filter out meaningless information or noises in input data while performing

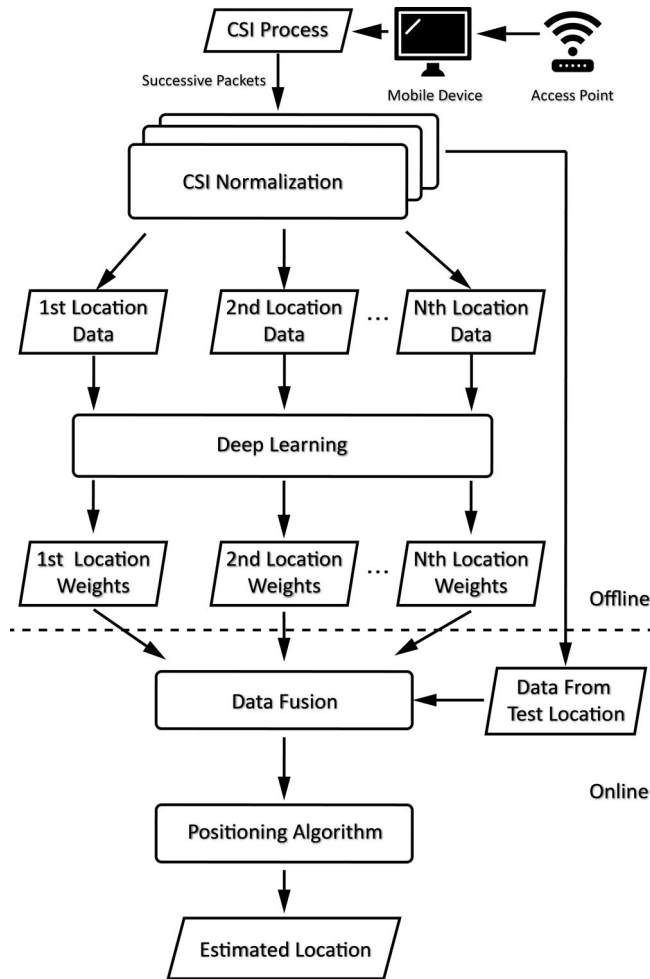


Figure 13. The DeepFi system. The normalized CSI data are separated according to their location. The deep learning method uses the CSI information to reconstruct the WiFi fingerprints with the weights of all the locations. Probabilistic methods are then used to estimate the user's location given a test CSI.

dimension reduction and then outputs the location weights of each position. In the prediction phase, the systems use probabilistic methods to get the final estimation.

Here, the classic DeepFi system (X. Wang, Gao, Mao, & Pandey, 2015, 2016) is considered. The DeepFi system is shown in Figure 13 and has main two phases: the off-line phase and the on-line phase. In the off-line phase, WiFi signals are received by a mobile device from APs and preprocessed before being passed to the deep learning methods. All the data are divided into several groups according to their location where the data were collected. The deep learning method (i.e. AE) detailed in Figure 14 generates the unique weight for each location as the position's new fingerprint. In the on-line phase, real-time test data is compared with the new fingerprint of each location. Probabilistic methods are then used to estimate the user's location. In the testbeds of a living room and a laboratory room, the MSEs of the DeepFi are 0.95 m and 1.8 m, respectively.

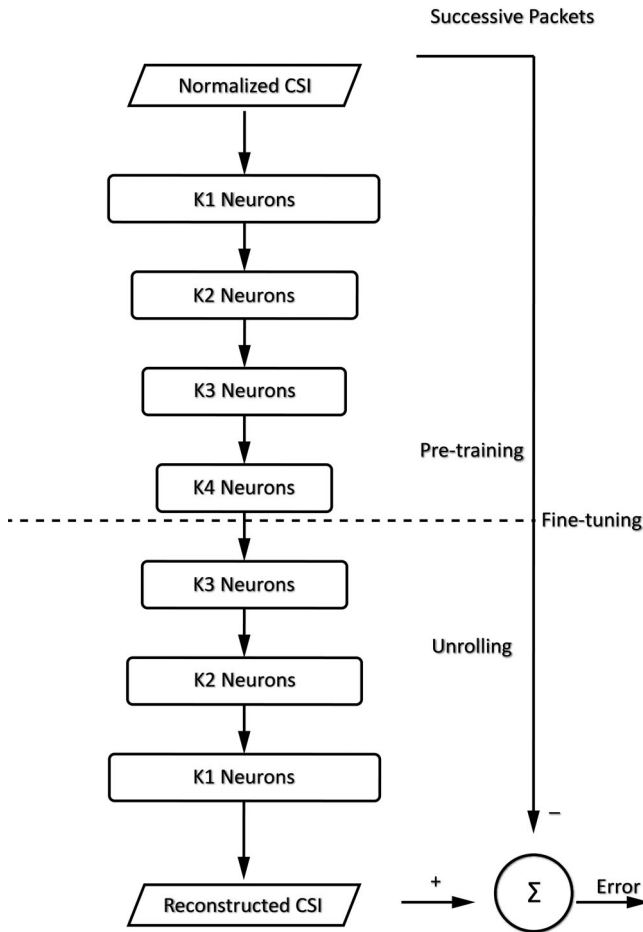


Figure 14. The neural network structure in DeepFi. The encoder part of the neural network labelled as ‘Pre-training’ in the figure is used to perform noise reduction and dimension reduction of the input data. K1, K2, K3 and K4 denote the number of neurons in the first, second, third, and fourth hidden layer, respectively.

In addition, X. Wang, Gao, and Mao (2015, 2016) adopt the CSI phase information as the input to the system. The CSI phase data is extracted by linear transformation. The weights of each position are trained by a greedy learning strategy. As introduced in these papers, the sub-network between two consecutive layers forms a restricted Boltzmann machine. X. Wang, Gao, and Mao (2017) proposes a system named Bi-loc which utilizes the bi-modal data (i.e. estimated angle of arrivals and average amplitudes derived from CSI data) to perform positioning estimation.

6.4. Positioning estimation algorithms

After hierarchical features are extracted and candidate positions are generated by neural networks, the final estimation of the user’s position need to be made. Since deep learning methods are a form of machine learning, the covered systems can carry on to perform the predictions.

The statistical result of the covered systems shows that 13 papers in total perform classification tasks which is predicting the floor the user is on while all of them are using majority voting as their prediction algorithm. Majority voting takes the classification results from the neural networks and choose the class that receives the largest number of classifications(or votes) as the final result. It could be defined as

$$C(X) = \text{mode}\{h_1(X), h_2(X), \dots, h_B(X)\} \quad (6)$$

where X is input WiFi data, $C(X)$ is the classification result of the X , *mode* represents the majority voting algorithm, $h_1(X), h_2(X), \dots, h_B(X)$ are the B number of classification results generated by the neural networks.

There are 56 systems that perform regression tasks in this section, 39.3% (22 out of 56) of them leverage probabilistic methods. They estimate the final position of the user by calculating the weighted centroid/weighted average of the candidate positions. And 12 of these systems adopt Bayes' law to calculate the posteriori probability which is used as the weight. Note that training positions are used as labels of the training WiFi data for final weighted average calculation and the predicted coordinates x and y are generated simultaneously under the regression circumstances. The final estimation of user's coordinates is calculated by the weighted average. The basic calculation of the weighted average is define as

$$(\hat{x}, \hat{y}) = \frac{\sum_{i=1}^N \alpha_i(x_i, y_i)}{\sum_{i=1}^N \alpha_i} \quad (7)$$

where $(\hat{x}, \hat{y})$ is the final estimated position, x_i and y_i are the coordinates of the i th candidate position, N represents the total number of the candidate positions, α_i is the corresponding weight of the i th candidate position.

Furthermore, 12 systems utilize Bayes' law to calculate the posteriori probabilities of all candidate positions in the training data. The weighted average is computed as follows:

$$(\hat{x}, \hat{y}) = \sum_{j=1}^K P(x_j, y_j) | D) \cdot (x_j, y_j) \quad (8)$$

$$P(x_j, y_j) | D) = \frac{P((x_j, y_j))P(D | (x_j, y_j))}{\sum_{k=1}^K P((x_k, y_k))P(D | (x_k, y_k))} \quad (9)$$

where $(\hat{x}, \hat{y})$ is the final estimated position, K is the total number of training positions, $P((x_j, y_j) | D)$ is the posteriori probability of (x_j, y_j) , D represents the training data, (x_j, y_j) is the j th training position, $P((x_j, y_j))$ and $P((x_k, y_k))$ are the prior probabilities of the j th and k th training position, respectively, $P(D | (x_j, y_j))$ and $P(D | (x_k, y_k))$ are the likelihood functions.

In addition, 21.4% (12 out of 56) of the regression systems adopt the popular machine learning method K-Nearest Neighbours (KNN) to make the final positioning estimation. After generating the new features of a new position based on the WiFi data collected by the user, the KNN algorithm measures the distance from the features of this new position to all training positions. The distance measure used can be Euclidean, Manhattan,

Minkowski or Weighted distance. Then the KNN selects the top K training positions closest to the new position and calculates their average as the final position estimation. Weighted K-Nearest Neighbours (WKNN) is the weighted version of KNN and is able to be more robust against variations in distances of the KNN which may lead to wrong decisions. Especially, the weight in WKNN could be the prediction probability of each selected training position to be the exact position where the user is. Algorithms used by the covered systems also include Extended Kalman Filter (EKF), Maximum Likelihood Estimation (MLE), dynamic Markov Decision Process (MDP), and Support Vector Regression (SVR). The main trend here is to calculate the weighted average of the candidate positions generated by deep learning methods.

6.5. Performance comparisons of systems employing deep learning as a feature extraction method

In this sub-section, Cumulative Distribution Function (CDF) plots and boxplots will be used to illustrate the performance comparisons of WiFi indoor positioning systems using deep learning methods to extract features.

The indoor positioning systems will be divided into 2 groups, systems regarding the positioning problem as a classification problem and systems regarding it as a regression problem. In the classification groups, all systems will be compared based on floor hitting rate and zone hitting rate, while in the regression group, systems will be compared based on their MDE and RMSE. There are 12 classification systems and 49 regression systems covered in this section. Note that some systems perform both types of positioning tasks which will be included in both groups.

Firstly, a detailed comparison will be made among classification systems. Due to the lack of diversity in the covered papers, there are only 13 classification systems proposed by researchers and all of them reported floor-level accuracy. All of these papers are based on WiFi RSS signals. Among them, only one system uses CNN, one uses AE, two use hybrid methods combining ANN and AE while the remaining systems apply ANN to extract features from the input data. The mean floor hitting rate of these indoor positioning systems is 92.59% and goes up to 93.97% after filtering out the outliers. In particular, Qi et al. (2018) achieved the best performance of 98.69% (the exact accuracy of this system is derived from the chart in their paper but not directly provided by the researchers) while the second best is Campos et al. (2014) with the floor hitting rate of 97%. Both papers took advantage of ANN neural network as the feature extraction method while Qi et al. (2018) employed ensemble ELM and Campos et al. (2014) used ANN. These two papers utilized the majority voting strategy as the final prediction algorithm because several individual ANN networks were used and their results were combined.

For the proposed regression systems covered in this section, their performances will be investigated and comparisons will be made according to the different evaluation metrics they used.

Due to the same reason as in the classification group, the number of papers using RMSE in the regression group is comparatively small. All of them are only based on WiFi RSS signals except Y. Li et al. (2019) uses hybrid signals from multiple sensors other than RSS. Specifically, Y. Li et al. (2019) adopts RSS to estimate the position uncertainty and further combines these results and the dead-reckoning (DR) solutions from

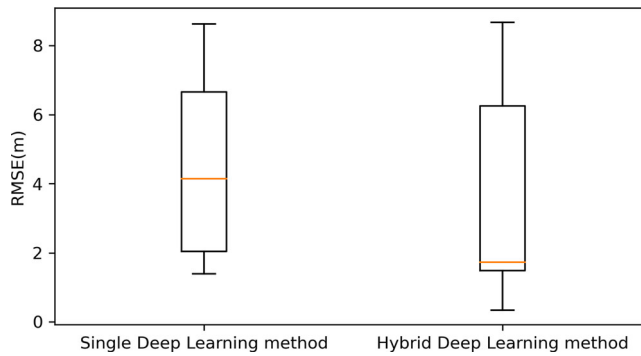


Figure 15. The boxplot shows RMSE results of the systems that use deep learning as feature extraction methods. It can be seen from the figure that utilizing more than one neural network could improve the RMSE results for a WiFi indoor positioning system. For instance, the best RMSE of 0.339 m is given by W. Zhang et al. (2016) which uses the combination of DNN and SDAE to extract features from the input data.

inertial sensors to estimate the user's position via Extended Kalman Filter (EKF). A diversity of neural networks were tried, 8 out of the 15 systems use ANN, 2 use DBN and the others utilize hybrid deep learning methods including the combination of ANN and SDAE, CNN and RNN and MDN, ANN and CNN and AE, AE and ANN.

For the sake of comparison, the papers are divided into two groups: one using a single neural network and the other using hybrid deep learning methods and adopting more than 1 neural network for feature extraction. The boxplot of their performance measured in RMSE is depicted in Figure 15. W. Zhang et al. (2016) gives the best performance as its RMSE for positioning is 0.339 m. W. Zhang et al. (2016) uses the hybrid methods of combining DNN and SDAE where SDAE is for the pre-training of DNN and then utilizes coarse localizer and HMM-based fine localizer for final estimation. The deep learning method here is for feature extraction and feature classification. It is surprising that the RMSE of an RSS-based indoor positioning system could achieve sub-metre level accuracy. The second best is Soro and Lee (2018) which uses multiple ANNs to achieve RMSE of 1.39 m. The mean RMSE of all covered papers is 4.18 m.

The number of papers using MDE as the evaluation metric reaches 50 which offers an opportunity to make a more comprehensive comparison. The performance of systems employing different deep learning feature extraction methods will be compared and the effect of using different WiFi signals as input will be investigated.

The general results are demonstrated by a CDF plot comparing the systems which have the top regression performances (see Figure 16). System 1 (proposed by T. Li et al., 2018) and system 2 (proposed by X. Wang, Gao, Mao, & Pandey, 2015) are based on CSI signals while system 3 (proposed by Hoang et al., 2019) and system 4 (proposed by Xue et al., 2020) are based on WiFi RSS. As illustrated in the CDF plot, systems based on CSI could produce less than 2 m distance error more than 98% of the time. On the other hand, systems utilizing RSS produce less than 3 m distance error more than 98% of the time. These results reveal that systems using CSI signals could achieve more stable and accurate positioning performances than those using RSS. Although RSS-based systems might get a distance error less than 0.5 m, 50% of

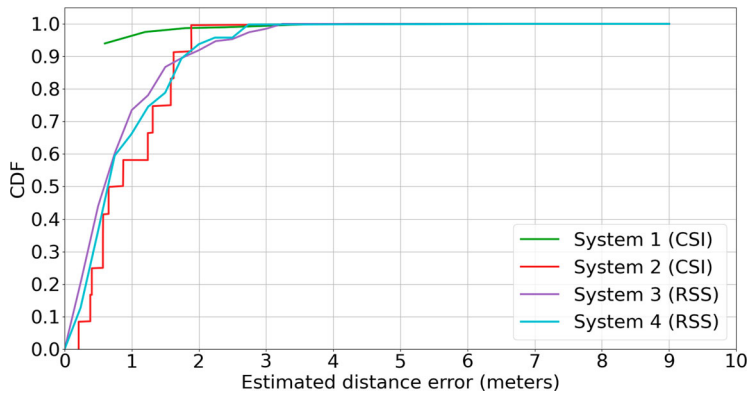


Figure 16. Comparison of the indoor positioning systems based on different WiFi signals. CSI-based systems could perform better, more stable and more accurate positioning estimation than RSS-based ones.

the time, their estimation deviations are comparatively large which leads to less reliable performances than those of CSI-based systems.

To investigate the positioning results further, a boxplot of MDE for different types of input is shown in Figure 17 where the CSI-based indoor positioning systems could have more stable predictions compared to the RSS-based systems. It is worth noticing that the median value of MDE for the RSS-based systems is lower than the CSI-based ones, the mean MDE of the RSS-based systems is 1.96 m while the CSI-based systems could achieve the mean MDE of 1.43 m. Furthermore, there are six positioning results generated by the systems based on CSI image and their mean value of MDE is also 1.43 m.

Next, the MDE performances of systems employing different neural networks are compared, see Figure 18. It is clear that both ANN and CNN perform generally better than other neural networks. The mean error of all indoor positioning systems using ANN is 1.607 m while that of CNN is 1.343 m. After filtering out the outliers, the mean MDE of

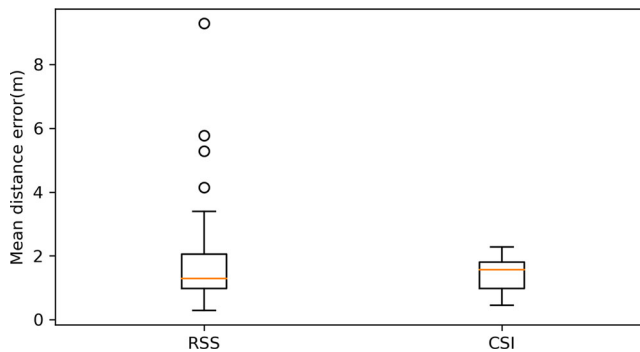


Figure 17. The boxplot of the MDE results for covered systems based on different inputs. Though WiFi RSS-based systems using deep learning as feature extraction methods could achieve generally better results than those based on CSI, the variation of their MDEs is larger. It is worth noticing that the best three RSS-based systems all utilize signals from other sensors (e.g. Inertial Measurement Unit (IMU)) to improve their positioning stability.

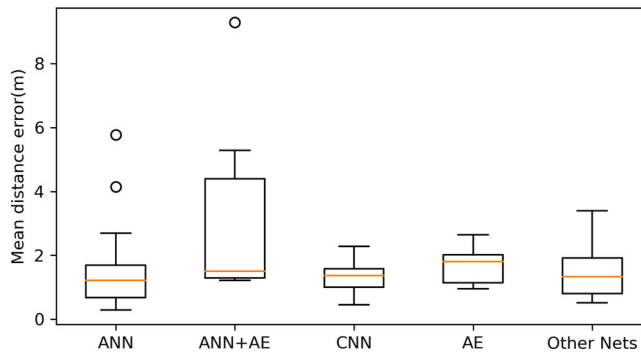


Figure 18. The MDE boxplot shows the effect of using different neural networks to extract features. Both ANN and CNN perform generally better than other neural networks. The mean MDE of indoor positioning systems using ANN is 1.607 m while that of CNN is 1.343 m. It is demonstrated in the boxplot that ANN, as a feature extraction method with less computational cost compared to CNN, is able to effectively generate meaningful information from the input data.

ANN could achieve 1.188 m. The average MDE of the group ‘Other Nets’ (i.e. systems using neural networks other than ANN, CNN and AE) is 1.555 m. Among these networks, systems using LSTM Hoang et al. (2019) and Capsule Network Own et al. (2019) both achieve sub-metre level accuracy. CNN and capsule network are the types of neural networks that extract information using a specific layer called the convolution layer from two-dimensional data such as images. Data transformed into two dimensions can also be fed into such networks. These networks are able to find higher-level information from the data since they convolute the 2D input data several times and use the condensed output as the evidence to perform location estimation.

Though it is not common for indoor positioning systems to achieve sub-metre level accuracy only using WiFi RSS, certain systems could still get very astonishing results by carefully modified models, for example Belmonte-Hernández et al. (2019), Own et al. (2019), G. Zhang et al. (2019), Xue et al. (2020), Hoang et al. (2019), Soro and Lee (2019) and D. V. Nguyen et al. (2018). The results of these systems are summarised in Table 3. In particular, it is not difficult to get sub-metre level accuracy for systems using information from multiple sensors other than RSS. For instance, Xingli et al. (2018) achieves the MDE of 0.29 m and Belmonte-Hernández et al. (2019) achieves 0.45 m. Using multiple sensors like iBeacon, Bluetooth Low Energy (BLE), Inertial Measurement Unit (IMU), and magnetometre can definitely enhance the performance. Detailed information about these sensors can be found in K. A. Nguyen et al. (2021) and therefore will not be included in this review. Furthermore, Zhao et al. (2019) and T. Li et al. (2018) which only use CSI as the input and achieve sub-metre level accuracy where Zhao et al. (2019) using CNN gets an MDE of 0.46 m and T. Li et al. (2018) using DNN gets an MDE of 0.6 m.

6.6. Trends and lessons learned in using deep learning as a feature extraction method

WiFi indoor positioning systems using deep learning as the feature extraction methods could produce stable prediction results. In this section, research papers tend to use RSS

Table 3. The RSS-based systems that achieve sub-metre level MDE.

Research work	Input	Network	MDE (m)
Own et al. (2019)	RSS	Capsule Network	0.99
G. Zhang et al. (2019)	RSS	CNN	0.9554
Xue et al. (2020)	RSS	ANN	0.87
Hoang et al. (2019)	RSS	LSTM	0.75
Soro and Lee (2019)	RSS	ANN	0.68
D. V. Nguyen et al. (2018)	RSS	ANN	0.62

more especially for floor prediction. And using multiple ANNs with the majority voting strategy could achieve the best floor prediction performance among all peer papers. For the regression indoor positioning systems, 32 out of 50 results are from RSS-based systems while there are only 18 CSI-based positioning results. It seems that RSS is still more popular in this area due to its easy accessibility. To achieve more stable and accurate positioning estimation, CSI is clearly the better choice. However, certain systems using RSS could also achieve sub-metre level accuracy if well modified deep learning methods were used to extract features. It is interesting to notice that RSS-based systems utilizing multiple sensors could achieve even better performances than CSI-based systems. Such results clearly imply that a combination of different signals from multiple sensors could greatly improve the performance of RSS-based indoor positioning systems. Considering that the sensors like IMU and magnetometre are common integrated sensors on smartphones, systems based on multi-sensors (e.g. Y. B. Bai et al., 2016; Xingli et al., 2018) could be easier to implement and popularize on smartphones in the near future. In terms of the neural network types, ANN is efficient enough for feature extraction. On the other hand, CNN is a better choice for WiFi indoor positioning systems aiming for more promising and accurate positioning estimation due to its convolution layer that could better extract hierarchical features from 2D data. If planning to use CNN as the extraction method, the input data needs to be modified as 2-dimensional data and the computation of CNN is more expensive than that of ANN. These facts make it harder to implement CNN than ANN.

7. Deep learning-based WiFi indoor positioning solutions

This section reviews those WiFi indoor positioning systems that use deep learning directly to predict the user's or object's location. Unlike the previous section, these systems will only be compared using positioning efficiency evaluation metrics including hitting rate, MDE, RMSE and CDF.

The basic structure of the systems reviewed in this section is illustrated in Figure 19. For any systems using deep learning directly as positioning solutions, the input data are also preprocessed before used by the positioning algorithms. The deep learning methods are utilized to perform location estimation which is the focus of this section. The deep learning neural networks included in this section could be classified as Artificial Neural Networks (ANN), Convolution Neural Network (CNN), Auto Encoder (AE), Deep Belief Network (DBN), Recurrent neural network (RNN) and other networks. Multilayer Perceptron (MLP), Deep Neural Network (DNN), Back Propagation Neural Network (BPNN), Extreme Learning Machine (ELM), Feed-forward Neural Network (FFNN), Discriminant-

adaptive Neural Network (DANN), etc are all grouped as ANN. Stacked Autoencoder (SAE) and Stacked Denoising Autoencoder (SDAE) are classified as AE while Long Short-Term Memory (LSTM) is considered as RNN.

Table 4 compares the WiFi indoor positioning systems that use deep learning directly for the estimation about the user's or object's location where the evaluation metrics are the MDE, RMSE, hitting rate and CDF. All the information about the systems are derived from the papers directly and the missing entries indicate that the papers did not provide the specific information. A more detailed comparison will be discussed in Section 7.4.

7.1. ANN

Koike-Akino et al. (2020) presented a system using a comparatively different WiFi signal – spatial beam Signal-to-Noise Ratios (SNRs). It is a medium-grained measurement of WiFi compared to the fine-grained CSI and coarse-grained RSS. The system uses the beam SNRs to form the fingerprinting database and implements ANN (a modified ResNet) on the database as shown in Figure 20. The input is the beam SNRs containing enriched information about spatial propagation paths of millimetre wave (mmWave) WiFi signals used during beam training phase. The beam SNRs are then passed to 3 residual blocks where there are shortcuts from the input to the output to maintain the residual gradient. The combination of beam SNRs and residual blocks ensures the improvement in the stability and accuracy of the WiFi indoor positioning system. Three main tasks: location classification, location-and-orientation classification, and coordinates prediction. The testbed consists of 6 offices and 8 cubicles filled with furniture during busy hours. 3 APs are fixed in a specific direction at fixed positions along the aisle. The data collection was conducted in 7 locations. The system achieves 100% correct location predictions and 99% of simultaneous location-and-orientation classification accuracy. The mean RMSE of the system reaches 11.1 cm for a direct coordinates estimation.

Xingli et al. (2018) considered the indoor positioning task as a regression problem. To enhance the coordinates prediction accuracy, the fusion of multi-source data is taken into consideration. The authors used the geomagnetic data, iBeacon signals and WiFi RSS signals in their database. All the multi-source data are passed to an RBM-initialized ANN (DNN to be specific). Other than using cross validation and grid search to fine-tune the neural network, Kalman Filter (KF) is also utilized to smooth the preprocessed data to simplify the input and retain important information. The testbed is two connected rooms of 124 m². One clear trajectory of user's movement is chosen with 15 reference points along the path and 1300 groups of data were collected in each position. As a result, the mean distance error of the system using both DNN and Kalman filter is 0.29 m, the maximum position error is 1.59 m, and the position error is within 1 m with the probability of 96.33%. As a comparison, the best MSE produced by other machine learning methods on the same testbed is 1.26 m by Quadratic Discriminant Analysis. It is astonishing for an RSS system to achieve such a small MDE. It seems that using multi-source data in some specific testbeds could largely enhance the regression accuracy of an indoor positioning system.

Table 4. Comparison of the covered WiFi-based indoor positioning systems using Deep Learning as a prediction method.

Research work	Input	Deep learning method	MSE (m)	RMSE (m)	Hitting rate (%)
J. Bai et al. (2021)	RSS	RNN+AE	1.6	–	–
Khassanov et al. (2021)	RSS	RNN LSTM	3.05	–	–
Khatab et al. (2021)	RSS	AE ELM	–	–	Zone:86.68
Ssekidde et al. (2021)	RSS	CNN	–	–	Room: 97.3; Point: 94.93
Alitaleshi et al. (2020)	RSS	ELM	–	–	Floor 98.1343
Careem et al. (2020)	RSS	MLP	–	–	Location: 94.38
Guney et al. (2020)	RSS	DNN	–	–	Location: 96.36
Nabati et al. (2020)	RSS	GAN MLP	–	–	Room: 97.2
Rizk et al. (2020)	RSS	DNN	2.62, 90% chance	–	–
Sun et al. (2020)	RSS	LSTM	4.29	–	–
S. Bai et al. (2019)	RSS	RNN	3.99	–	–
BelMannoubi and Touati (2019)	RSS	SAE	–	–	Room: 85.58
Z. Chen et al. (2019)	RSS	LSTM	1.615	–	–
Chidlovskii and Antsfeld (2019)	RSS	VAE	–	4.65	–
Elbes et al. (2019)	RSS	LSTM	1.01	–	–
Lembo et al. (2019)	RSS	ANN	0.41, 90% chance	–	–
Z. Liu et al. (2019)	RSS	CNN	3.91	–	Location: 84.17 (15days after)
Lian et al. (2019)	RSS	ELM	1, 90% chance	–	–
Song, Fan, Xiang, et al. (2019)	RSS	CNN+SAE	7.6	–	Building: 100; Floor: 96.03
Song, Fan, He, et al. (2019)	RSS	CNN+SAE	–	–	Building: 100; Floor: 95.92
Turabieh and Sheta (2019)	RSS	Recurrent Neural Network	–	Building: 96.45; Floor: 87.41; Office room: 95.80	–
Turgut et al. (2019)	RSS	AE	–	–	Grid: 95.95
R. Wang et al. (2019)	RSS	SDAE+MLP	4.25	–	–
H. Wang et al. (2019)	RSS	ELM	–	2.93	–
Adege, Yayeh, et al. (2018)	RSS	ANN	–	0.56	–
Adege, Yen, et al. (2018)	RSS	FFNN+MLP	–	0.32	Room: 100
Aikawa et al. (2018)	RSS	ANN	18	–	Location: 87
Anzum et al. (2018)	RSS	CPN	–	–	Zone: 91
Cheerla and Ratnam (2018)	RSS	PMNN	–	–	–
De Vita and Bruneo (2018)	RSS	DNN	–	–	Room: dynamic 83.6; static 99.6
H. Y. Hsieh et al. (2018)	RSS	RNN & LSTM	2.482	–	Floor: 99.7
Research work	Input	Deep learning method	MSE (m)	RMSE (m)	Hitting rate (%)
Ibrahim et al. (2018)	RSS	CNN	2.77	–	Building and Floor: 100
Jang and Hong (2018)	RSS	CNN	–	–	Building and Floor: 95.41
Joseph and Sasi (2018)	RSS	DNN SAE	–	–	Floor: 93

(Continued)

Table 4. Continued.

Research work	Input	Deep learning method	MSE (m)	RMSE (m)	Hitting rate (%)
Kim (2018)	RSS	DNN SDAE	7.93	–	Floor: 94.15
Kim, Wang, et al. (2018)	RSS	DNN SAE	–	–	Real scenario: building 99; floor: 93.429; location: 97.198. UJIIndoorloc: location <70%
Lin et al. (2018)	RSS	CNN	1.105	–	–
Park et al. (2018)	RSS	DNN	6.943	–	Floor: 94.40
T. Zhang and Yi (2018)	RSS	CNN	–	–	Location: 90.83
Zhong et al. (2018)	RSS	SAE	–	–	Floor/Area: 94
Zhu et al. (2018)	RSS	GAN	–	1, 80% chance	–
Belay et al. (2017)	RSS	FFNN	0.4, 90% chance	–	–
Basiouny et al. (2017)	RSS	SMN	1.85	–	–
Lukito and Chrismanto (2017)	RSS	RNN	–	–	Room: 82.47
Nowicki and Wietrzykowski (2017)	RSS	DNN SAE	–	–	Building and Floor: 92
Xiao et al. (2017)	RSS	DNN	7.6	–	–
H. Dai et al. (2016)	RSS	ANN AE	0.75	–	–
Félix et al. (2016)	RSS	DNN DBN	1.01	–	–
He et al. (2016)	RSS	DBN	–	–	Area: 96.5
N. Li et al. (2016)	RSS	ANN	1.893	–	–
C. Zhou and Wieser (2016)	RSS	BPNN ANN	4, 90% chance	–	–
Gu et al. (2015)	RSS	ELM	7.5	–	Grid: 34 40
Lu et al. (2014)	RSS	ELM	–	2.95	–
Fahed and Liu (2013)	RSS	DNN	1.895	–	–
Mok and Cheung (2013)	RSS	FFNN	1 to 4	2.045	–
Mehmood et al. (2010)	RSS	ANN	1.43	–	–
Borenović and Nešković (2009)	RSS	ANN	8.9 absolute MSE	–	–
Borenovic et al. (2008)	RSS	ANN	9.58 absolute MSE	Room type: 62.34; Room: 11.2	–
Ding et al. (2008)	RSS	BPNN/MLP	4.42	–	Floor: 98.15
Fang and Lin (2008)	RSS	DANN	<2	–	–
Tsai et al. (2008)	RSS	BPN	2, 92% chance	–	–
Stella et al. (2007)	RSS	MLP	1.79	–	–
Vilović and Zovko-Cihlar (2005)	RSS	MLP	0.2 absolute MSE	–	–
Battiti et al. (2002)	RSS	MLP	2.3	–	–
Research work	Input	Deep learning method	MSE (m)	RMSE (m)	Hitting rate (%)
Roy and Chowdhury (2021)	RSS+Wi-FiID	Ensemble Learning	2.4	–	Zone: 94.21
Adege, Lin, et al. (2018b)	RSS+Wi-FiID	DNN	<0.97	–	Room: 100

(Continued)

Table 4. Continued.

Research work	Input	Deep learning method	MSE (m)	RMSE (m)	Hitting rate (%)
Adege, Lin, et al. (2018a)	RSS+WiFiID	DNN MLP	–	0.69	Zone: 96.64
Gan et al. (2017)	RSS+WiFiID	DBN	0.52	–	–
Jiang et al. (2018)	RSS of WiFi & BLE	ELM	3, 95.5% chance	–	–
Farid et al. (2016)	RSS of WiFi & WSN	ANN	1.05	–	–
Y. Liu et al. (2020)	RSS image	CNN	2, 46.76% chance	–	–
Sinha and Hwang (2019)	RSS image	CNN	1.44	–	Location: 94.95
W. Shao, Luo, Zhao, Ma, et al. (2018)	RSS image +magnetic filed	CNN MLP	1, 90% chance	–	–
W. Zhang et al. (2017)	RSS+magnetic filed	DNN SAE	1.4551	–	–
Y. B. Bai et al. (2016)	RSS+magnetic filed	ANN	–	<2.6	–
Xingli et al. (2018)	RSS+iBeacon +Geomagnetic	DNN RBM	0.29	–	–
H. Zhang et al. (2019)	RSS CSI	DNN	6.448	–	–
C. H. Hsieh et al. (2019)	RSS/CSI amplitude	CNN	0.92, 99.97% chance	–	–
W. Liu et al. (2020)	CSI amplitude	DNN	0.78	–	–
Z. Zhang et al. (2020)	CSI amplitude	1D-CNN	1.2669	–	–
F. Wang et al. (2019)	CSI amplitude	CNN	–	–	Location: 95.68
Berruet et al. (2018)	CSI amplitude	CNN	1.39, 50%	–	–
Cai et al. (2018)	CSI amplitude	CNN	1.843	–	Location: office 94.55; corridor 96.40
R. Zhou et al. (2018)	CSI amplitude	DNN	1.29	–	–
G. Wang et al. (2021b)	CSI phase	1D-CNN +LSTM	1.825	–	Location: 88
G. Wang et al. (2021a)	CSI phase	1D+CNN +LSTM	1.33	–	Grid: 88.4
Dang et al. (2020)	CSI phase	DHNN	1.6	–	–
C. Xu et al. (2016)	CSI phase	AE	0.58, 80% chance	–	–
Koike-Akino et al. (2020)	beam SNR	DNN(ResNet)	0.111	Location: 100; Location +Orientations: 98.96	–
Y. Xu et al. (2009)	SNR	ANFIS	2.77	–	–
Ohta et al. (2015)	WiFiID	ANN	–	–	Location: 91.2

7.2. CNN

Sinha and Hwang (2019) proposed a system based on CNN using the image representation of WiFi RSS signals to perform indoor positioning. The main idea is to convert the RSS signals to 2D images and process the images using the neural network. 74 reference points were set in the testbed where RSS signals from 256 APs were collected. These 256 RSS measurements collected in a single reference point are transformed to a 16×16 image as shown in Figure 21 where the light dots indicate that

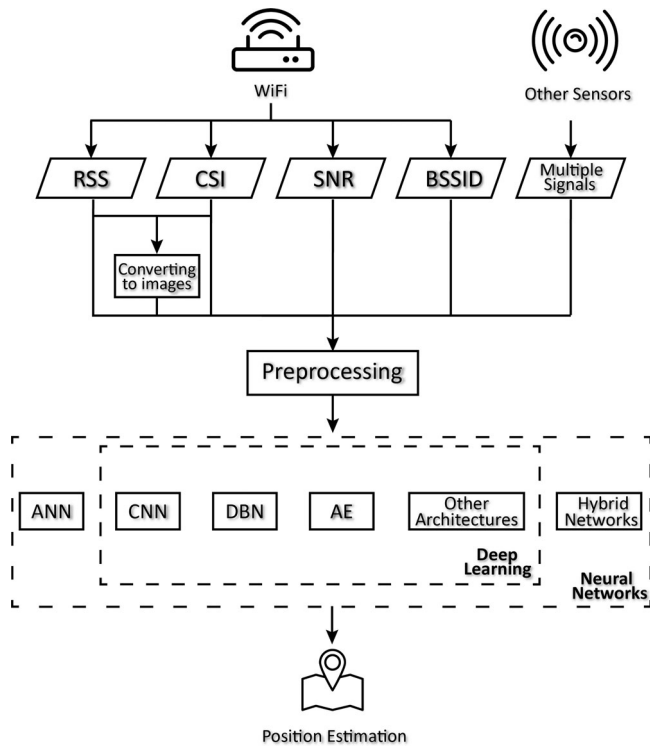


Figure 19. The general process of WiFi indoor positioning systems employing deep learning as prediction methods. Please note that the difference between deep learning and neural networks is identified in this figure. While in the comparisons, all the different neural networks adopted by covered systems will still be compared whether they are deep neural networks or simple neural networks.

the RSS values from these APs can be received at the current reference point. During the preprocessing, the input RSS data are enriched by simply facilitating augmentation and utilizing mean values and uniform random numbers to add information into the dataset (see Sinha & Hwang, 2019). Then a six-layer neural network was proposed to predict which reference point the user is on as shown in Figure 22. Compared to the existing CNN models such as AlexNet, ResNet, ZFNet, Inception v3, and MobileNet v2, the proposed system achieves the better positioning accuracy of 94.45% and an MSE of 1.44 m.

C. H. Hsieh et al. (2019) compared several combinations of different neural networks (1D-CNN and MLP) and different input (CSI and RSS), and discussed the best system based on CSI while using 1D-CNN as prediction method. The architecture of the neural network is presented in Figure 23. This 1D-CNN is different from general CNN because the convolutional layers in 1D-CNN only have 1-dimensional small filters and deal with 1D data rather than 2D image data. Since the original CSI data is 1D data, it is appropriate to use such 1D-CNN which ensures the accuracy of the system and reduces high computational costs. The extracted information derived from CSI is utilized to determine the user location. The testbed is a room of 13.82 m × 8.58 m filled with obstacles. A total of 251,388 CSI measurements were

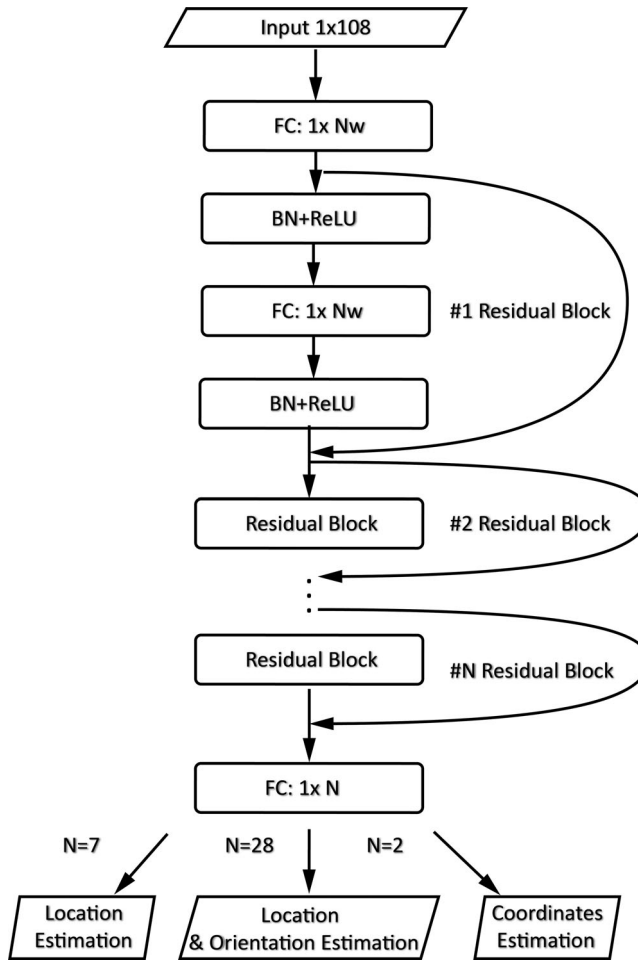


Figure 20. The architecture of the neural network proposed by Koike-Akino et al. (2020). ‘BN’ represents batch normalization which is a deep learning approach to normalize the data. This network utilizes the residual blocks to maintain the residual gradient from the input data and perform location-only classification, simultaneous location-and-orientation classification and direct coordinates estimation based on different output layers.

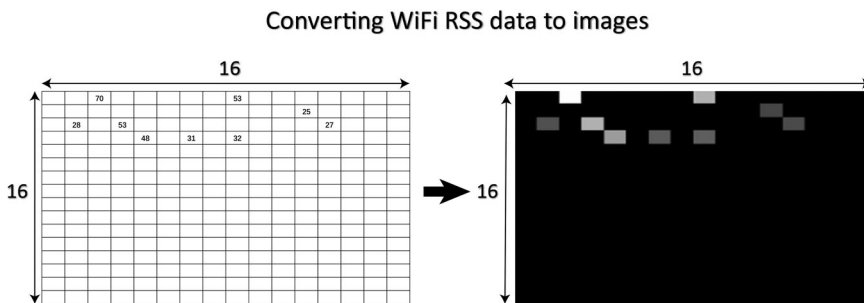


Figure 21. The transformation from the RSS signals to a 2D-image data. Each 16×16 greyscale image is converted from 256 RSS signals. The brightness of the pixel in the image represents how strong the corresponding RSS is. The black dots in the image indicate that RSS from the corresponding APs can not be received.

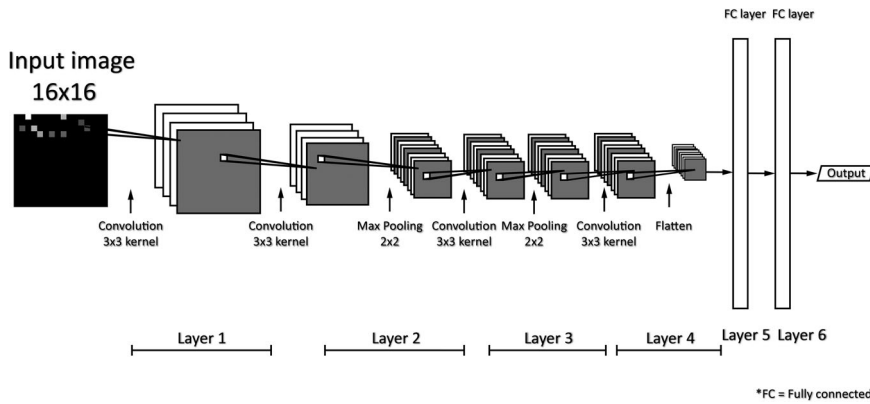


Figure 22. The architecture of the CNN in Sinha and Hwang (2019). This neural network contains 4 convolutional layers and fully connected layers to perform classification.

collected of which 90% were used for training and the rest for testing. The authors divided the room into 16 blocks and used the system to predict the exact block where the user is standing in. To study the robustness of the system, 3 testers of different body shapes were participating in the experiment. The results showed that the system based on 1D-CNN and CSI data reaches the maximum error of 0.92 m with the probability of 99.97%. However, further validation of the system using large public datasets is needed due to the small size of the testbed.

7.3. AE

Kim, Wang, et al. (2018) proposed a system based on stacked autoencoder (SAE) and ANN to estimate the building and floor the user is on. The system considers position estimation as multi-class classification ones. However, it does not predict a single sample's targeted building and floor level at the same time. Instead, it estimates the building, floor and

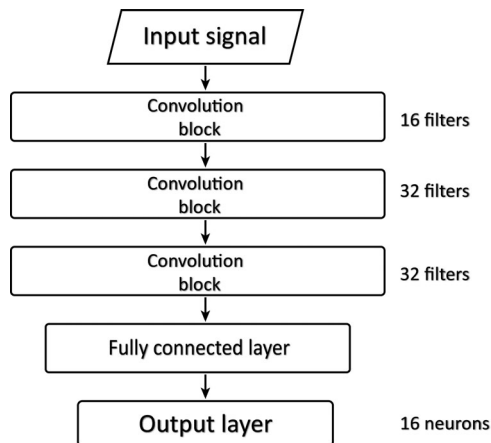


Figure 23. The architecture of the 1D-CNN in C. H. Hsieh et al. (2019).

location separately as shown in [Figure 24](#). Thus, multiple classifiers are adopted by the indoor positioning system. The input RSS data are processed by the SAE for dimension reduction and noise filtering. After preprocessing, the hierarchical features of the RSS data are fed into different classifiers for building, floor and location predictions. For the building and floor predictions, this system achieves 99% for building hitting rate and 93.429% for floor hitting rate. For the floor-level location estimation, the testbed is the fourth floor of the EE building in Xi'an Jiaotong-Liverpool University (XJTLU) campus. Around 200 APs were set in the environment where more than 4000 RSS fingerprints were collected. The floor-level location estimation accuracy reaches 97.198%. However, the floor-level location accuracy goes below 70% when applying the same system to the public WiFi dataset UJIIndoorLoc. The authors assumed this is due to the much larger number of locations and the closeness of the corresponding fingerprints in the public dataset. Since AE is mostly used to reduce the dimension complexity and noises of the data, it remains a challenge for researchers to improve the performance of AE for indoor positioning in a complex environment.

7.4. Performance comparisons of systems employing deep learning as a positioning estimation method

In this sub-section, performance comparisons are made among WiFi indoor positioning systems using deep learning methods directly to predict the user's location. The indoor positioning systems will be divided into 2 groups, systems regarding the positioning problem as a regression problem and systems regarding it as a classification problem. In the regression group, all systems will be compared based on their MDE and RMSE, while in the classification group, floor hitting rate and zone hitting rate will be used. There are 42 classification systems and 58 regression systems considered here. It is worth noticing that there are some systems performing both types of positioning problems which will be compared in both groups. Classification algorithms will be analysed first and then a detailed comparison will be made among the regression algorithms.

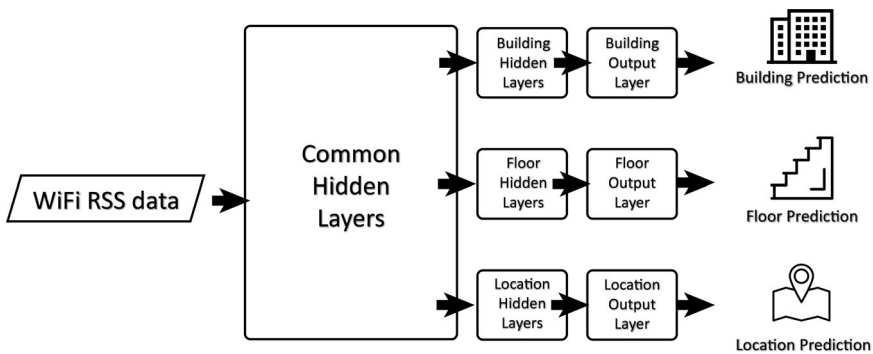


Figure 24. The structure of the DNN proposed by Kim, Wang, et al. (2018). The system treats the multi-label classification questions as multi-class classification ones. Since it predicts the building, floor and location via different output layers, the system becomes scalable and flexible and could be implemented easily in different indoor positioning scenarios.

For floor-level indoor positioning systems, RSS is the common input type of all covered papers. Thus, the main focus will be on different neural networks. According to the main types of neural networks used, the systems are divided into 5 groups. They are systems using ANN, AE, Hybrid AE, DBN and other networks (i.e. CNN, RNN, LSTM), and their average floor hitting rates are 90.81%, 93.60%, 94.89%, 94.45% and 93.40%, respectively (see Figure 25). It can be seen from Figure 25 that ANN and DBN perform comparatively better in floor-level prediction though their variances are relatively higher. In fact, among the top 5 floor prediction systems with floor hitting rates above 98%, 3 of them apply ANN (Alitalishi et al., 2020; Ding et al., 2008; He et al., 2016), one system (He et al., 2016) uses DBN, and one system (H. Y. Hsieh et al., 2018) uses LSTM.

A more accurate positioning method is to locate the user to a preset grid or position. As mentioned before, different researchers used different terms in their testbeds. In order to compare the systems fairly, the term 'Zone' is used to describe the targeted class or location of these systems whether it is a grid, a block or an area. However, the size of the zone used in different systems vary from $1\text{ m} \times 1\text{ m}$ to a single room. Thus, the results represented in the boxplot may not be able to offer a comprehensive view of all zone-predicting systems. It can be seen from Figure 26 that CSI offers better and more stable performance than RSS signals or even hybrid RSS input signals (i.e. a combination of WiFi RSS and other sensor measurements) in the zone prediction setting. The overall mean zone hitting rate of all covered papers is 83.44%, while the mean zone hitting rate of the hybrid RSS-based systems is 77.89%, 82.59% for the RSS-based systems and 91.83% for the CSI-based systems.

The performance comparison of zone predicting systems using different neural networks is presented in Figure 27. The average zone hitting rate for ANN-based systems is 79.16%, for CNN-based systems is 83.43%, for systems using other neural networks (e.g. AE, RNN, and counter propagation network (CPN)) is 91.18% and for using hybrid networks (e.g. GAN+ANN, ANN+AE, or CNN+LSTM) is 89.19%. However, the average zone hitting rate will reach 93.26% if the outliers are filtered out for the CNN-based systems.

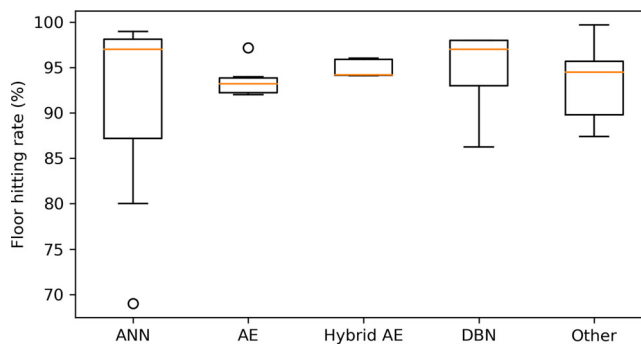


Figure 25. The boxplot shows the floor hitting rate results for systems using deep learning as a prediction method. The papers are grouped according to the main types of neural networks they use. It is illustrated in the boxplot that ANN and DBN are better in floor-level prediction while the variances in both groups are comparatively higher. Among the top 5 floor prediction systems with floor hitting rate above 98%, 3 of them apply ANN.

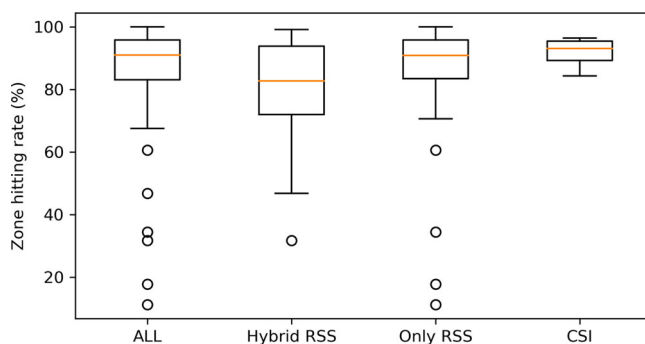


Figure 26. The boxplot shows the zone hitting rate results for systems using deep learning as prediction method. This boxplot mainly compares the effect of using different WiFi signals as the input. CSI is more stable than RSS signals or even hybrid RSS input signals. The hybrid RSS input signals mean a combination of RSS and signals from other sensors (e.g. magnetometre, accelerometer, and Bluetooth).

It can be seen from the boxplot that CNN is a generally better choice for the zone prediction task.

Among the top 3 systems with the zone hitting rate above 98% shown in Table 5, all of them apply ANN to perform zone prediction and use RSS or SNR as the input. Note that SNR is the signal-to-noise ratios of the WiFi signals. But only two papers use SNR (i.e. Koike-Akino et al., 2020; Y. Xu et al., 2009) which makes it hard to evaluate such signal measure as a separate input type and draw a convincing conclusion. The best results gained by using CNN is Ssekidde et al. (2021) with the zone hitting rate of 97.3%.

For regression systems, the first evaluation metric used is RMSE. The number of papers using RMSE is relatively small which leads to no further division of the covered papers. The mean RMSE of all 9 papers is 2.39 m while the best result is from Koike-Akino et al. (2020) with 0.095 m and 0.111 m RMSE using millimetre wave WiFi and a modified DNN based on the famous residual neural network ResNet, respectively. Interestingly, Koike-Akino et al.

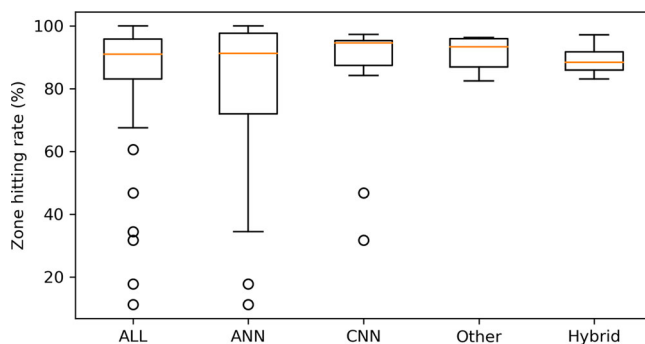


Figure 27. The boxplot shows the zone hitting rate results for systems using deep learning as a prediction method. This boxplot mainly compares the effect of using different neural networks. Though some systems based on ANN could achieve the best result in the comparison, the variance of all ANN systems is astonishingly large which represents the instability of ANN systems. It could be deduced by the boxplot that CNN is generally better in zone predicting.

Table 5. The top 3 zone predicting systems that utilize deep learning as prediction solution.

Research work	Input	Network	Zone hitting rate (%)
Adege, Yen, et al. (2018)	RSS	ANN	100
Koike-Akino et al. (2020)	SNR	ANN	100
De Vita and Bruneo (2018)	RSS	ANN	99.6

(2020) is the one using SNR as input data. The next 2 best systems (Adege, Lin, et al., 2018a, 2018b) have RMSE of 0.32 and 0.55 m. All these three papers use the ANN to perform regression. It could be concluded from such astonishing results that systems using ANN are able to produce very accurate positioning estimation. Moreover, systems using RSS could achieve sub-metre level RMSE. Given the surprising performance reported by Koike-Akino et al. (2020), it is worth considering SNR as the input, application of residual network or even utilizing mmWave WiFi device as APs.

As mentioned in the previous section, MDE is the most popular and most direct evaluation metric for WiFi-based regression indoor positioning systems. To give a detailed analysis of the results, all related systems will be compared based on different types of input and then a comparison will be made based on different neural networks adopted by the indoor positioning systems, see Figures 28–30.

It can be seen from Figure 28 that systems based on hybrid RSS signals achieve the best performance while systems based on CSI are the second best. The mean MDE of all related papers is 3.65 m and it goes down to 3.45 m after filtering out the outliers. The average MDE of all systems using RSS only as the input is 4.06 m. The average MDE of hybrid RSS-based systems is 4.03 m while that of the CSI-based systems is 1.85 m. After the removal of outliers, the mean MDE of hybrid RSS-based systems is reduced to 1.25 m and that of the CSI-based system group is 1.47 m. As a result, CSI could provide more accurate and stable prediction about the user's location as expected and the hybrid RSS signals could greatly improve the positioning accuracy of the RSS-based systems.

The large number of published CSI-based systems enables us to analyse the effect of utilizing CSI amplitude or phase as the input. As illustrated in Figure 29, CSI amplitude-based systems have better performances than CSI phase-based ones. The average MDE

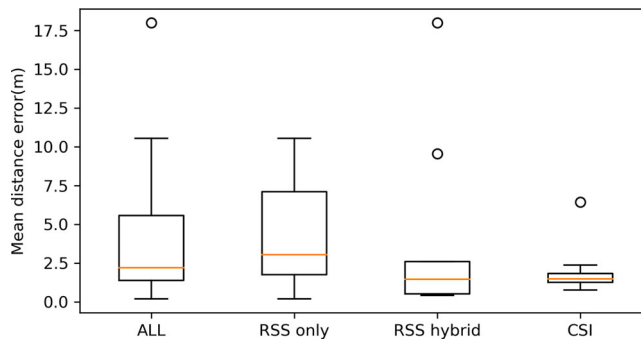


Figure 28. The boxplot of MDE results from systems using deep learning as the prediction method. This boxplot aims to compare the effect of using different WiFi signals as the input to the positioning system. CSI could provide more accurate and stable results about the user's location, while hybrid RSS signals could largely improve the RSS-based positioning accuracy.

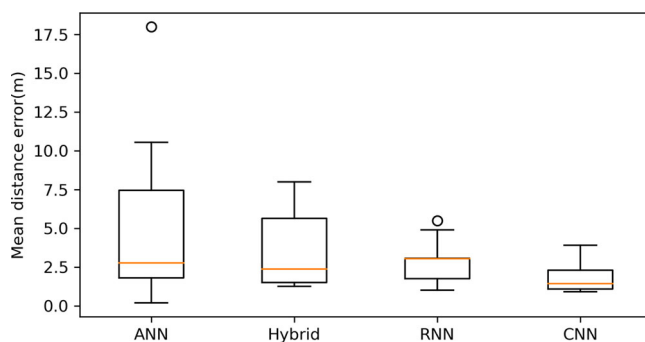


Figure 30. The boxplot compares the MDE from systems using different neural networks. Due to the comparatively good performance and relatively higher number of systems using CNN, it could be derived from the figure that CNN is the best neural network to perform coordinates prediction.

is 1.23 m for systems using CSI amplitude and 1.71 m for those using CSI phase. It is demonstrated in the comparison that among all CSI-based indoor positioning systems, those employing the CSI amplitude could get a more accurate position estimation.

The comparison of MDE from all related indoor positioning systems is illustrated in Figure 30 where the reviewed papers are classified into 4 groups based on the major types of neural networks they used. They are systems using ANN, hybrid Neural Networks (i.e. CNN+SAE, ANN+SDAE, RNN+AE, CNN+ANN and CNN+LSTM), RNN, and CNN. The boxplot shows that CNN-based indoor positioning systems are the best. The overall mean MDE of all systems is 3.242 m. The average MDE for the systems based on hybrid networks, RNN, CNN are 3.690, 2.674 and 1.871 m, respectively. The mean MDE for the systems using ANN is 4.73 and 4.34 m after filtering out the outliers. Note that the mean MDE for AE-based systems is only 0.92 m. However, these results are only from two research papers which means lack of representativeness. Thus, CNN is the best neural network for coordinates prediction (i.e. predicting the user's exact position in a Cartesian coordinate system) because of its comparatively good performances and relatively higher number of reported systems.

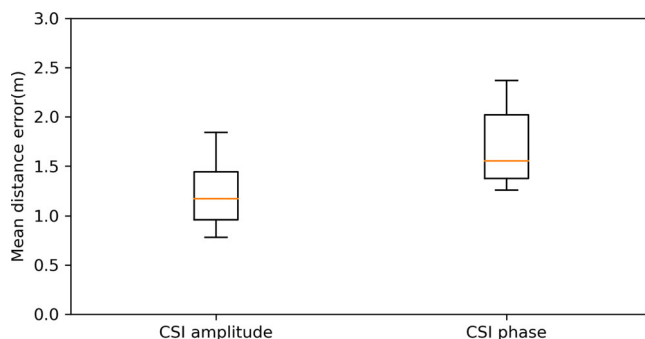


Figure 29. The effect of using CSI amplitude or CSI phase as the input to the neural network. Either for the general performance or the variance in the results, CSI amplitude is a better choice for indoor positioning systems to get more accurate position estimation.

Table 6. All systems reach the sub-metre level MDE while utilizing deep learning as prediction method.

Research work	Input	Network	MDE (m)
Vilović and Zovko-Cihlar (2005)	RSS	ANN	0.20
H. Dai et al. (2016)	RSS	AE	0.46
Adege, Lin, et al. (2018b)	hybrid RSS	ANN	0.50
H. Dai et al. (2016)	RSS	AE	0.67
W. Liu et al. (2020)	CSI amplitude	ANN	0.78
C. H. Hsieh et al. (2019)	CSI amplitude	CNN	0.92
Lin et al. (2018)	RSS	CNN	0.95

When looking at the best systems with sub-metre level MDE shown in Table 6, Vilović and Zovko-Cihlar (2005) has a mean absolute error of 0.2 m which is surprisingly good given it was published in 2005 and only using RSS as the input and MLP as the prediction algorithm. Among all indoor positioning systems that achieve sub-metre level MDE, there are mainly 3 types of input data which are hybrid RSS, RSS, and CSI amplitude. And for the neural networks, ANN, CNN, AE, DBN and RNN are the major types here. Due to that the testbeds used by these systems are different, the comparison result is less conclusive. In fact, there is no clear trend on how to get the best performance for the coordinates prediction, which signal type a system should use as the input and which neural network as the prediction algorithm. But it is clear that an RSS-based indoor positioning system could have a chance to achieve sub-metre level accuracy, either using a carefully modified neural network or utilizing hybrid signals from different sensors.

7.5. Trends and lessons learned in using deep learning as a positioning estimation method

Several trends can be identified from all WiFi indoor positioning systems using deep learning to estimate the user's location. For the classification systems, ANN and DBN are comparatively better in floor-level prediction. Moreover, ANNs are the best neural networks among all the covered zone hitting classification systems because the best performance is achieved by ANN-based indoor positioning systems. But generally speaking, researchers may consider using CNN as the prediction method for the zone prediction problem. In terms of the input type, CSI is generally the best WiFi signal which a system can choose to get promising zone hitting accuracy. However, the performance of systems using SNR could not be ignored. Due to lack of conclusive evidence, the effectiveness of SNR is yet to be explored. RSS if used well with appropriately modified deep learning models or coupled with other sensor signals, can also serve as a good choice of input for zone prediction systems.

For regression systems, systems using ANN are able to get very good positioning results. And systems using RSS could also have sub-metre level RMSE with properly modified deep learning models. Given the surprising performance of Koike-Akino et al. (2020), it is worth considering SNR as the input or using residual network or even using mmWave WiFi device as the access point. Based on the evaluation metric of MDE, systems using CSI could get more accurate and stable estimation about the user's location as expected. RSS-based systems using hybrid signals can greatly improve their positioning accuracy. Our comparison shows CNN is the best neural network to perform coordinates

prediction. The large number of systems utilizing CNN in the reviewed papers support such trend. Moreover, the performances of systems employing AE are promising. There is no clear result to identify which neural network architecture or which input type shall be used by the positioning system to get the best performance. However, it could be concluded that RSS-based indoor positioning systems could have a chance to achieve sub-metre level accuracy by either using a carefully modified neural network or utilizing hybrid signals from different sensors.

8. Conclusion and future perspectives

This article has reviewed more than 150 related research papers employing deep learning for WiFi indoor systems. These papers have been divided into two different categories. The first one employs deep learning approaches as feature extraction methods for WiFi indoor positioning. The second one utilizes deep learning models as regressors or classifiers when dealing with WiFi signals. Within each category, the systems are compared individually and then the effect of applying different neural networks and inputs are analysed. A set of performance metrics are devised to evaluate such systems, including hitting rate, Mean Distance Error, Root Mean Squared Error, Cumulative Distribution Function, Complexity (i.e. the number of hidden layers) and testing time.

To answer the research question *'What is the most accurate WiFi signal measure for indoor positioning systems?'*, our review indicates that RSS is still a comparatively useful technology that could be employed as the system input to perform positioning estimation, while CSI achieved much more accurate estimation on average. Among all systems utilizing RSS, some could achieve sub-metre level accuracy while the majority of them reach metre-level accuracy. Using signals from other sensors like IMU and magnetometre could significantly improve the positioning performance of the RSS-based systems to sub-metre level. The CSI signals could provide more stable and abundant information for sub-metre level positioning, but the hardware limitation makes it more challenging to adopt. Thus, combining WiFi RSS signals and signals from multiple sensors could be a potential research direction in WiFi indoor positioning. The newly released WiFi RTT technology is an attractive option to achieve sub-metre level accuracy. However, there has not been sufficient results in the literature to draw a conclusion yet.

For the research question *'What are the most efficient neural networks for WiFi-based indoor positioning systems?'*, neural networks like ANN and DBN are simpler with acceptable performance accuracy, while CNN is much more complex with higher demand for computational resources but achieves better results. Considering the majority of future indoor positioning users rely more and more on smartphones, it is wise to seek for WiFi indoor positioning systems that could be easily implemented on such devices. It is worth noticing that although smartphones have an aggregation of multiple sensors, their computational ability is limited. Hence, proposing a system based on RSS and signals from multiple sensors while employing simple neural networks (e.g. ANN and DBN) could be a good option in the near future. If the system is going to perform the location estimation on-line with a remote server, then these limitations could be alleviated. In that case, CNN with its best feature extracting capability should be the optimal solution in future WiFi indoor positioning systems.

Acknowledgments

We would like to thank the reviewers for their insightful and constructive comments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research is funded by University of Brighton's Connected Futures, Radical Futures' initiatives, and the School of Computing, Engineering and Mathematics' QR grant.

Notes on contributors

Xu Feng is studying for his doctor's degree at University of Brighton. Xu holds a bachelor's degree of Engineering at Zhejiang University. His main research interest is in indoor positioning.

Dr. Khuong An Nguyen holds a Ph.D, an M.Phil (Cantab), and a B.Sc (Hons) at Royal Holloway University and University of Cambridge. Dr. Nguyen is a Fellow of the Centre for Machine Learning at Royal Holloway, and an Associate Fellow of the Royal Institute of Navigation. His research interests cover epidemic contact tracing and healthcare monitoring.

Dr. Zhiyuan Luo is currently a Professor in the Department of Computer Science and a member of Centre for Machine Learning at Royal Holloway, University of London, UK. His main research interests are in machine learning, data analysis, networked systems and agent-based computing, and applications of these algorithms and techniques. His research has been funded by UK Engineering and Physical Science Research Council (EPSRC), Medical Research Council (MRC), Royal Society and European Union Framework 7 and Horizon 2020 programme.

ORCID

Xu Feng  <http://orcid.org/0000-0003-2181-6220>

Khuong An Nguyen  <http://orcid.org/0000-0001-6198-9295>

Zhiyuan Luo  <http://orcid.org/0000-0002-3336-3751>

References

- Abbas, M., Elhamshary, M., Rizk, H., Torki, M., & Youssef, M. (2019). WiDeep: WiFi-based accurate and robust indoor localization system using deep learning. In *2019 IEEE International Conference on Pervasive Computing and Communications (PERCOM)* (pp. 1–10), IEEE.
- Adege, A. B., Lin, H. P., G. B. Tarekegn, & Jeng, S. S. (2018a). Applying deep neural network (DNN) for robust indoor localization in multi-building environment. *Applied Sciences*, 8(7), 1062. <https://doi.org/10.3390/app8071062>
- Adege, A. B., Lin, H. P., Tarekegn, G. B., Munaye, Y. Y., & Yen, L. (2018b). An indoor and outdoor positioning using a hybrid of support vector machine and deep neural network algorithms. *The Journal of Sensors*, 2018(1), 1–12. <https://doi.org/10.1155/2018/1253752>
- Adege, A. B., Yayeh, Y., Berie, G., Lin, H. P., Yen, L., & Li, Y. R. (2018). Indoor localization using K-nearest neighbor and artificial neural network back propagation algorithms. In *2018 27th Wireless and Optical Communication Conference (WOCC)* (pp. 1–2), IEEE.
- Adege, A. B., Yen, L., Lin, H. P., Yayeh, Y., Li, Y. R., Jeng, S. S., & Berie, G. (2018). Applying deep neural network (DNN) for large-scale indoor localization using feed-forward neural network (FFNN)

- algorithm. In *2018 IEEE International Conference on Applied System Invention (ICASI)* (pp. 814–817), IEEE.
- Aikawa, S., Yamamoto, S., & Morimoto, M. (2018). WLAN finger print localization using deep learning. In *2018 IEEE Asia-Pacific Conference on Antennas and Propagation (APCAP)* (pp. 541–542), IEEE.
- Alitalishi, A., Jazayeriy, H., & Kazemitabar, S. J. (2020). WiFi fingerprinting based floor detection with hierarchical extreme learning machine. In *2020 10th International Conference on Computer and Knowledge Engineering (ICCKE)* (pp. 113–117), IEEE.
- Anzum, N., Afroze, S. F., & Rahman, A. (2018). Zone-based indoor localization using neural networks: A view from a real testbed. In *2018 IEEE International Conference on Communications (ICC)* (pp. 1–7), IEEE.
- Bai, J., Sun, Y., Meng, W., & Li, C. (2021). Wi-Fi fingerprint-based indoor mobile user localization using deep learning. *Wireless Communications and Mobile Computing*, 2021(1), 1–12. <https://doi.org/10.1155/2021/6660990>
- Bai, S., Yan, M., Wan, Q., He, L., Wang, X., & Li, J. (2019). DL-RNN: An accurate indoor localization method via double RNNs. *IEEE Sensors Journal*, 20(1), 286–295. <https://doi.org/10.1109/JSEN.7361>
- Bai, Y. B., Gu, T., & Hu, A. (2016). Integrating Wi-Fi and magnetic field for fingerprinting based indoor positioning system. In *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 1–6), IEEE.
- Basiouny, Y., Arafa, M., & Sarhan, A. M. (2017). Enhancing Wi-Fi fingerprinting for indoor positioning system using single multiplicative neuron and PCA algorithm. In *2017 12th International Conference on Computer Engineering and Systems (ICCES)* (pp. 295–305), IEEE.
- Basri, C., & El Khadimi, A. (2016). Survey on indoor localization system and recent advances of WIFI fingerprinting technique. In *2016 5th International Conference on Multimedia Computing and Systems (ICMCS)* (pp. 253–259), IEEE.
- Battiti, R., Le, N. T., & Villani, A. (2002). Location-Aware Computing: A Neural Network Model For Determining Location In Wireless LANs. In *2002 IEEE Int. Semicond. Conf.*, IEEE.
- Belay, A., Yen, L., Renu, S., Lin, H. P., & Jeng, S. S. (2017). Indoor localization at 5 GHz using dynamic machine learning approach (DMLA). In *2017 International Conference on Applied System Innovation (ICASI)* (pp. 1763–1766), IEEE.
- BelMannoubi, S., & Touati, H. (2019). Deep neural networks for indoor localization using WiFi fingerprints. In *International Conference on Mobile, Secure, and Programmable Networking* (pp. 247–258), Springer.
- Belmonte-Hernández, A., Hernández-Peñaloza, G., D. M. Gutiérrez, & Alvarez, F. (2019). SWiBluX: Multi-sensor deep learning fingerprint for precise real-time indoor tracking. *IEEE Sensors Journal*, 19(9), 3473–3486. <https://doi.org/10.1109/JSEN.7361>
- Bernas, M., & Placzek, B. (2015). Fully connected neural networks ensemble with signal strength clustering for indoor localization in wireless sensor networks. *International Journal of Distributed Sensor Networks*, 11(12), 403242. <https://doi.org/10.1155/2015/403242>
- Berruet, B., Baala, O., Caminada, A., & Guillet, V. (2018). DelFin: A deep learning based CSI fingerprinting indoor localization in IoT context. In *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 1–8), IEEE.
- Borenovic, M., Neskovic, A., Budimir, D., & Zeliz, L. (2008). Utilizing artificial neural networks for WLAN positioning. In *2008 IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications* (pp. 1–5), IEEE.
- Borenović, M. N., & Nešković, A. M. (2009). Positioning in WLAN environment by use of artificial neural networks and space partitioning. *Annals of Telecommunications-Annales des Télécommunications*, 64(9), 665–676. <https://doi.org/10.1007/s12243-009-0115-0>
- Cai, C., Deng, L., Zheng, M., & Li, S. (2018). PILC: Passive indoor localization based on convolutional neural networks. In *2018 Ubiquitous Positioning, Indoor Navigation and Location-based Services (UPINLBS)* (pp. 1–6), IEEE.
- Campos, R. S., Lovisolo, L., & de Campos, M. L. R. (2014). Wi-Fi multi-floor indoor positioning considering architectural aspects and controlled computational complexity. *Expert Systems with Applications*, 41(14), 6211–6223. <https://doi.org/10.1016/j.eswa.2014.04.011>

- Careem, A. A., Ali, W. H., & Jasim, M. H. (2020). Wirelessly Indoor Positioning System based on RSS Signal. In *2020 International Conference on Computer Science and Software Engineering (CSASE)* (pp. 238–243), IEEE.
- Cheerla, S., & Ratnam, D. V. (2018). RSS based Wi-Fi positioning method using multi layer neural networks. In *2018 Conference on Signal Processing and Communication Engineering Systems (SPACES)* (pp. 58–61), IEEE.
- Chen, H., Zhang, Y., Li, W., Tao, X., & Zhang, P. (2017). ConFi: Convolutional neural networks based indoor Wi-Fi localization using channel state information. *IEEE Access*, 5(1), 18066–18074. <https://doi.org/10.1109/ACCESS.2017.2749516>.
- Chen, Z., Zou, H., Yang, J., Jiang, H., & Xie, L. (2019). WiFi fingerprinting indoor localization using local feature-based deep LSTM. *IEEE Systems Journal*, 14(2), 3001–3010. <https://doi.org/10.1109/JSYST.4267003>
- Chidlovskii, B., & Antsfeld, L. (2019). Semi-supervised variational autoencoder for WiFi indoor localization. In *2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 1–8), IEEE.
- Chollet, F. (2018). *Deep learning with python*. (Vol. 361). Manning.
- Dai, H., Ying, W. h., & Xu, J. (2016). Multi-layer neural network for received signal strength-based indoor localisation. *IET Communications*, 10(6), 717–723. <https://doi.org/10.1049/cmu2.v10.6>
- Dai, P., Yang, Y., Wang, M., & Yan, R. (2019). Combination of DNN and improved KNN for indoor location fingerprinting. *Wireless Communications and Mobile Computing*, 2019(1), 1–9. <https://doi.org/10.1155/2019/4283857>.
- Dang, X., Tang, X., Hao, Z., & Ren, J. (2020). Discrete Hopfield neural network based indoor Wi-Fi localization using CSI. *EURASIP Journal on Wireless Communications and Networking*, 2020(1), 1–16. <https://doi.org/10.1186/s13638-020-01692-7>
- De Vita, F., & Bruneo, D. (2018). A deep learning approach for indoor user localization in smart environments. In *2018 IEEE International Conference on Smart Computing (SMARTCOMP)* (pp. 89–96), IEEE.
- Ding, X., Li, H., Li, F., & Wu, J. (2008). A novel infrastructure WLAN locating method based on neural network. In *Proceedings of the 4th Asian Conference on Internet Engineering* (pp. 47–55), Bangkok.
- Dinh-Van, N., Nashashibi, F., Thanh-Huong, N., & Castelli, E. (2017). Indoor Intelligent Vehicle localization using WiFi received signal strength indicator. In *2017 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)* (pp. 33–36), IEEE.
- Dou, F., Lu, J., Wang, Z., Xiao, X., Bi, J., & Huang, C. H. (2018). Top-down indoor localization with Wi-fi fingerprints using deep Q-network. In *2018 IEEE 15th International Conference on Mobile ad hoc and Sensor Systems (MASS)* (pp. 166–174), IEEE.
- Elbakly, R., Aly, H., & Youssef, M. (2018). TrueStory: Accurate and robust RF-based floor estimation for challenging indoor environments. *IEEE Sensors Journal*, 18(24), 10115–10124. <https://doi.org/10.1109/JSEN.2018.2872827>
- Elbakly, R., & Youssef, M. (2020). The StoryTeller: Scalable building-and ap-independent deep learning-based floor prediction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(1), 1–20. <https://doi.org/10.1145/3380979>
- Elbes, M., Almaita, E., Alrawashdeh, T., Kanan, T., AlZu'bi, S., & Hawashin, B.. (2019). An indoor localization approach based on deep learning for indoor location-based services. In *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)* (pp. 437–441), IEEE.
- Fahed, D., & Liu, R. (2013). Wi-Fi-based localization in dynamic indoor environment using a dynamic neural network. *International Journal of Machine Learning and Computing*, 3(1), 127. <https://doi.org/10.7763/IJMLC.2013.V3.286>
- Fang, S. H., & Lin, T. N. (2008). Indoor location system based on discriminant-adaptive neural network in IEEE 802.11 environments. *IEEE Transactions on Neural Networks*, 19(11), 1973–1978. <https://doi.org/10.1109/TNN.2008.2005494>
- Farid, Z., Nordin, R., Ismail, M., & Abdullah, N. F. (2016). Hybrid indoor-based WLAN-WSN localization scheme for improving accuracy based on artificial neural network. *Mobile Information Systems*, 2016(1), 6923931. <https://doi.org/10.1155/2016/6923931>.

- Félix, G., Siller, M., & Alvarez, E. N. (2016). A fingerprinting indoor localization algorithm based deep learning. In *2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN)* (pp. 1006–1011), IEEE.
- Gan, X., Yu, B., Huang, L., & Li, Y. (2017). Deep learning for weights training and indoor positioning using multi-sensor fingerprint. In *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 1–7), IEEE.
- Gu, Y., Chen, Y., Liu, J., & Jiang, X. (2015). Semi-supervised deep extreme learning machine for Wi-Fi based localization. *Neurocomputing*, 166(C), 282–293. <https://doi.org/10.1016/j.neucom.2015.04.011>.
- Guney, S., Erdogan, A., Aktas, M., & Ergun, M. (2020). Wi-Fi based indoor positioning system with using deep neural network. In *2020 43rd International Conference on Telecommunications and Signal Processing (TSP)* (pp. 225–228), IEEE.
- Haider, A., Wei, Y., Liu, S., & Hwang, S. H. (2019). Pre-and post-processing algorithms with deep learning classifier for Wi-Fi fingerprint-based indoor positioning. *Electronics*, 8(2), 195. <https://doi.org/10.3390/electronics8020195>
- He, S., & Chan, S. H. G. (2015). Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons. *IEEE Communications Surveys & Tutorials*, 18(1), 466–490. <https://doi.org/10.1109/COMST.2015.2464084>
- He, S., Tan, J., & Chan, S. H.G. (2016). Towards area classification for large-scale fingerprint-based system. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 232–243), ACM.
- Hinton, G. E. (2009). Deep belief networks. *Scholarpedia*, 4(5), 5947. <https://doi.org/10.4249/scholarpedia.5947>
- Hoang, M. T., Yuen, B., Dong, X., Lu, T., Westendorp, R., & Reddy, K. (2019). Recurrent neural networks for accurate RSSI indoor localization. *IEEE Internet of Things Journal*, 6(6), 10639–10651. <https://doi.org/10.1109/JIoT.6488907>
- Hsieh, C. H., Chen, J. Y., & Nien, B. H. (2019). Deep learning-based indoor localization using received signal strength and channel state information. *IEEE Access*, 7(1), 33256–33267. <https://doi.org/10.1109/ACCESS.2019.2903487>.
- Hsieh, H. Y., Prakosa, S. W., & Leu, J. S.. (2018). Towards the implementation of recurrent neural network schemes for WiFi fingerprint-based indoor positioning. In *2018 IEEE 88th Vehicular technology Conference (VTC-Fall)* (pp. 1–5), IEEE.
- Hsu, C. S., Chen, Y. S., Juang, T. Y., & Wu, Y. T. (2019). An adaptive Wi-Fi indoor localisation scheme using deep learning. *International Journal of Ad Hoc and Ubiquitous Computing*, 30(4), 265–274. <https://doi.org/10.1504/IJAHUC.2019.098880>
- Hu, X., Chu, L., Pei, J., Liu, W., & Bian, J. (2021). Model complexity of deep learning: A survey. Preprint arXiv:2103.05127.
- Ibrahim, M., Torki, M., & ElNainay, M. (2018). CNN based indoor localization using RSS time-series. In *2018 IEEE Symposium on Computers and Communications (ISCC)* (pp. 01044–01049), IEEE.
- Jang, J. W., & Hong, S. N. (2018). Indoor localization with WiFi fingerprinting using convolutional neural network. In *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)* (pp. 753–758), IEEE.
- Jiang, X., Chen, Y., Liu, J., Gu, Y., & Hu, L. (2018). FSELM: Fusion semi-supervised extreme learning machine for indoor localization with Wi-Fi and Bluetooth fingerprints. *Soft Computing*, 22(11), 3621–3635. <https://doi.org/10.1007/s00500-018-3171-4>
- Joseph, R., & Sasi, S. B. (2018). Indoor positioning using WiFi fingerprint. In *2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET)* (pp. 1–3), IEEE.
- JunLin, G., Xin, Z., HuaDeng, W., & Lan, Y. (2020). WiFi fingerprint positioning method based on fusion of autoencoder and stacking mode. In *2020 International Conference on Culture-Oriented Science & Technology (ICCST)* (pp. 356–361), IEEE.
- Khassanov, Y., Nurpeiissov, M., Sarkytbayev, A., Kuzdeuov, A., & Varol, H. A. (2021). Finer-level sequential WiFi-based indoor localization. In *2021 IEEE/SICE International Symposium on System Integration (SII)* (pp. 163–169), IEEE.

- Khatab, Z. E., Gazestani, A. H., Ghorashi, S. A., & Ghavami, M. (2021). A fingerprint technique for indoor localization using autoencoder based semi-supervised deep extreme learning machine. *Signal Processing*, 181(1), 107915. <https://doi.org/10.1016/j.sigpro.2020.107915>.
- Kim, K. S. (2018). Hybrid building/floor classification and location coordinates regression using a single-input and multi-output deep neural network for large-scale indoor localization based on Wi-Fi fingerprinting. In *2018 Sixth International Symposium on Computing and Networking Workshops (CANDARW)* (pp. 196–201), IEEE.
- Kim, K. S., Lee, S., & Huang, K. (2018). A scalable deep neural network architecture for multi-building and multi-floor indoor localization based on Wi-Fi fingerprinting. *Big Data Analytics*, 3(1), 1–17. <https://doi.org/10.1186/s41044-018-0031-2>
- Kim, K. S., Wang, R., Zhong, Z., Tan, Z., Song, H., Cha, J., & Lee, S. (2018). Large-scale location-aware services in access: Hierarchical building/floor classification and location estimation using Wi-Fi fingerprinting based on deep neural networks. *Fiber and Integrated Optics*, 37(5), 277–289. <https://doi.org/10.1080/01468030.2018.1467515>
- Koike-Akino, T., Wang, P., Pajovic, M., Sun, H., & Orlik, P. V. (2020). Fingerprinting-based indoor localization with commercial MMWave WiFi: A deep learning approach. *IEEE Access*, 8(1), 84879–84892. <https://doi.org/10.1109/Access.6287639>.
- Kozma, R., Alippi, C., Choe, Y., & Morabito, F. C. (2018). *Artificial intelligence in the age of neural networks and brain computing*. Academic Press.
- Laoudias, C., Kemppi, P., & Panayiotou, C. G. (2009). Localization using radial basis function networks and signal strength fingerprints in WLAN. In *Globecom 2009–2009 IEEE Global Telecommunications Conference* (pp. 1–6). IEEE.
- Le, D. V., Meratnia, N., & Havinga, P. J. (2018). Unsupervised deep feature learning to reduce the collection of fingerprints for indoor localization using deep belief networks. In *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 1–7), IEEE.
- Lembo, S., Horsmanheimo, S., Somersalo, M., Laukkanen, M., Tuomimäki, L., & Huilla, S. (2019). Enhancing WiFi RSS fingerprint positioning accuracy: Lobe-forming in radiation pattern enabled by an air-gap. In *2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 1–8), IEEE.
- Li, H., Zeng, X., Li, Y., Zhou, S., & Wang, J. (2019). Convolutional neural networks based indoor Wi-Fi localization with a novel kind of CSI images. *China Communications*, 16(9), 250–260. <https://doi.org/10.1109/CC.6245522>
- Li, J., Li, Y., & Ji, X. (2016). A novel method of Wi-Fi indoor localization based on channel state information. In *2016 8th International Conference on Wireless Communications & Signal Processing (WCSP)* (pp. 1–5), IEEE.
- Li, N., Chen, J., Yuan, Y., Tian, X., Han, Y., & Xia, M. (2016). A Wi-Fi indoor localization strategy using particle swarm optimization based artificial neural networks. *International Journal of Distributed Sensor Networks*, 12(3), 4583147. <https://doi.org/10.1155/2016/4583147>
- Li, T., Wang, H., Shao, Y., & Niu, Q. (2018). Channel state information–based multi-level fingerprinting for indoor localization with deep learning. *International Journal of Distributed Sensor Networks*, 14(10), 1550147718806719. <https://doi.org/10.1177/1550147718806719>
- Li, Y., Gao, Z., He, Z., Zhuang, Y., Radi, A., Chen, R., & El-Sheimy, N. (2019). Wireless fingerprinting uncertainty prediction based on machine learning. *Sensors*, 19(2), 324. <https://doi.org/10.3390/s19020324>
- Lian, L., Xia, S., Zhang, S., Wu, Q., & Jing, C. (2019). Improved Indoor positioning algorithm using KPCA and ELM. In *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)* (pp. 1–5), IEEE.
- Lin, W. Y., Huang, C. C., Duc, N. T., & Manh, H. N. (2018). Wi-Fi indoor localization based on multi-task deep learning. In *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)* (pp. 1–5), IEEE.
- Liu, C., Wang, C., & Luo, J. (2020). Large-scale deep learning framework on FPGA for fingerprint-based indoor localization. *IEEE Access*, 8(1), 65609–65617. <https://doi.org/10.1109/Access.6287639>.

- Liu, J., Liu, N., Pan, Z., & You, X. (2018). AutLoc: Deep autoencoder for indoor localization with RSS fingerprinting. In *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)* (pp. 1–6), IEEE.
- Liu, M., Chen, R., Li, D., Chen, Y., Guo, G., Cao, Z., & Pan, Y. (2017). Scene recognition for indoor localization using a multi-sensor fusion approach. *Sensors*, 17(12), 2847. <https://doi.org/10.3390/s17122847>
- Liu, W., Chen, H., Deng, Z., Zheng, X., Fu, X., & Cheng, Q. (2020). LC-DNN: Local connection based deep neural network for indoor localization with CSI. *IEEE Access*, 8(1), 108720–108730. <https://doi.org/10.1109/Access.6287639>.
- Liu, Y., Sinha, R. S., Liu, S. Z., & Hwang, S. H. (2020). Side-information-aided preprocessing scheme for deep-learning classifier in fingerprint-based indoor positioning. *Electronics*, 9(6), 982. <https://doi.org/10.3390/electronics9060982>
- Liu, Z., Dai, B., Wan, X., & Li, X. (2019). Hybrid wireless fingerprint indoor localization method based on a convolutional neural network. *Sensors*, 19(20), 4597. <https://doi.org/10.3390/s19204597>
- Lu, X., Long, Y., Zou, H., Yu, C., & Xie, L. (2014). Robust extreme learning machine for regression problems with its application to WiFi based indoor positioning system. In *2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)* (pp. 1–6), IEEE.
- Lukito, Y., & Chrisanto, A. R. (2017). Recurrent neural networks model for WiFi-based indoor positioning system. In *2017 International Conference on Smart Cities, Automation & Intelligent Computing Systems (ICON-SONICS)* (pp. 121–125), IEEE.
- Ma, Z., Wu, B., & Poslad, S. (2019). A WiFi RSSI ranking fingerprint positioning system and its application to indoor activities of daily living recognition. *International Journal of Distributed Sensor Networks*, 15(4), 1550147719837916. <https://doi.org/10.1177/1550147719837916>
- Mehmood, H., Tripathi, N. K., & Tipdecho, T. (2010). Indoor positioning system using artificial neural network. *Journal of Computer Science*, 6(10), 1219. <https://doi.org/10.3844/jcssp.2010.1219.1225>
- Mok, E., & Cheung, B. K. (2013). An improved neural network training algorithm for Wi-Fi fingerprinting positioning. *ISPRS International Journal of Geo-information*, 2(3), 854–868. <https://doi.org/10.3390/ijgi2030854>
- Nabati, M., Navidan, H., Shahbazian, R., Ghorashi, S. A., & Windridge, D. (2020). Using synthetic data to enhance the accuracy of fingerprint-based localization: A deep learning approach. *IEEE Sensors Letters*, 4(4), 1–4. <https://doi.org/10.1109/LSENS.7782634>
- Nguyen, D. V., De Charette, R., Nashashibi, F., Dao, T. K., & Castelli, E. (2018). WiFi fingerprinting localization for intelligent vehicles in car park. In *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 1–6), IEEE.
- Nguyen, D. V., Recalde, M. E. V., & Nashashibi, F. (2016). Low speed vehicle localization using WiFi fingerprinting. In *2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV)* (pp. 1–5), IEEE.
- Nguyen, K. A., Luo, Z., Li, G., & Watkins, C. (2021). A review of smartphones-based indoor positioning: Challenges and applications. *IET Cyber-Systems and Robotics*, 3(1), 1–30. <https://doi.org/10.1049/csy.v3.1>
- Nowicki, M., & Wietrzykowski, J. (2017). Low-effort place recognition with WiFi fingerprints using deep learning. In *International Conference Automation* (pp. 575–584), Springer.
- Ohta, M., Sasaki, J., Takahashi, S., & Yamashita, K. (2015). WiFi positioning system without AP locations for indoor evacuation guidance. In *2015 IEEE 4th Global Conference on Consumer Electronics (GCCE)* (pp. 483–484), IEEE.
- Own, C. M., Hou, J., & Tao, W. (2019). Signal fuse learning method with dual bands WiFi signal measurements in indoor positioning. *IEEE Access*, 7(1), 131805–131817. <https://doi.org/10.1109/Access.6287639>.
- Park, C. U., Shin, H. G., & Choi, Y. H. (2018). A parallel artificial neural network learning scheme based on radio wave fingerprint for indoor localization. In *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)* (pp. 794–797), IEEE.
- Qi, G., Jin, Y., & Yan, J. (2018). RSSI-based floor localization using principal component analysis and ensemble extreme learning machine technique. In *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)* (pp. 1–5), IEEE.

- Qian, W., Lauri, F., & Gechter, F. (2019). Convolutional mixture density recurrent neural network for predicting user location with WiFi fingerprints. Preprint arXiv:1911.09344.
- Rizk, H., Yamaguchi, H., Youssef, M., & Higashino, T. (2020). Gain without pain: Enabling fingerprinting-based indoor localization using tracking scanners. In *Proceedings of the 28th International Conference on Advances in Geographic Information Systems* (pp. 550–559), ACM.
- Roy, P., & Chowdhury, C. (2021). Designing an ensemble of classifiers for smartphone-based indoor localization irrespective of device configuration. *Multimedia Tools and Applications*, 80(1), 1–25. <https://doi.org/10.1007/s11042-020-10456-w>
- Shao, W., Luo, H., Zhao, F., Ma, Y., Zhao, Z., & Crivello, A. (2018). Indoor positioning based on fingerprint-image and deep learning. *IEEE Access*, 6(1), 74699–74712. <https://doi.org/10.1109/Access.6287639>.
- Shao, W., Luo, H., Zhao, F., Wang, C., Crivello, A., & Tunio, M. Z. (2018). DePos: Accurate orientation-free indoor positioning with deep convolutional neural networks. In *2018 Ubiquitous Positioning, Indoor Navigation and Location-based Services (UPINLBS)* (pp. 1–7), IEEE.
- Shao, Y., Li, L., & Guo, X. (2019). A semi-supervised deep learning approach towards localization of crowdsourced data. In *Proceedings of the ACM Turing Celebration Conference-China* (pp. 1–5), ACM.
- Sinha, R. S., & Hwang, S. H. (2019). Comparison of CNN applications for RSSI-based fingerprint indoor localization. *Electronics*, 8(9), 989. <https://doi.org/10.3390/electronics8090989>
- Song, X., Fan, X., He, X., Xiang, C., Ye, Q., Huang, X., Fang, G., Chen, L. L., Qin, J., & Wang, Z. (2019). CNNLoc: Deep-learning based indoor localization with WiFi fingerprinting. In *2019 IEEE Smartworld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)* (pp. 589–595), IEEE.
- Song, X., Fan, X., Xiang, C., Ye, Q., Liu, L., Wang, Z., He, X., Yang, N., & Fang, G. (2019). A novel convolutional neural network based indoor localization framework with WiFi fingerprinting. *IEEE Access*, 7(1), 110698–110709. <https://doi.org/10.1109/Access.6287639>.
- Soro, B., & Lee, C. (2018). Performance comparison of indoor fingerprinting techniques based on artificial neural network. In *TENCON 2018-2018 IEEE Region 10 Conference* (pp. 0056–0061), IEEE.
- Soro, B., & Lee, C. (2019). A wavelet scattering feature extraction approach for deep neural network based indoor fingerprinting localization. *Sensors*, 19(8), 1790. <https://doi.org/10.3390/s19081790>
- Ssekidde, P., Steven Eyobu, O., Han, D. S., & Oyana, T. J. (2021). Augmented CWT features for deep learning-based indoor localization using WiFi RSSI data. *Applied Sciences*, 11(4), 1806. <https://doi.org/10.3390/app11041806>
- Stella, M., Russo, M., & Begusic, D. (2007). Location determination in indoor environment based on RSS fingerprinting and artificial neural network. In *2007 9th International conference on Telecommunications* (pp. 301–306), IEEE.
- Sun, H., Zhu, X., Liu, Y., & Liu, W. (2020). WiFi based fingerprinting positioning based on Seq2seq model. *Sensors*, 20(13), 3767. <https://doi.org/10.3390/s20133767>
- Torres-Sospedra, J., Montoliu, R., Martínez-Usó, A., Avariento, J. P., Arnau, T. J., Benedito-Bordonau, M., & Huerta, J. (2014). UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems. In *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 261–270), IEEE.
- Tsai, C. Y., Chou, S. Y., Lin, S. W., & Wang, W. H.. (2008). Location determination of mobile device for indoor WLAN application using neural network. *Knowledge and Information Systems*, 20(1), 81–93. <https://doi.org/10.1007/s10115-008-0154-2>
- Turabieh, H., & Sheta, A. (2019). Cascaded layered recurrent neural network for indoor localization in wireless sensor networks. In *2019 2nd International Conference on New Trends in COMPUTING Sciences (ICTCS)* (pp. 1–6), IEEE.
- Turgut, Z., Üstebay, S., Aydın, G. Z. G., & Sertbaş, A. (2019). Deep learning in indoor localization using WiFi. In *International Telecommunications Conference* (pp. 101–110), Springer.
- Vilović, I., & Zovko-Cihlar, B. (2005). WLAN location determination model based on the artificial neural networks. In *Proceedings ELMAR-2005* (pp. 287–290), IEEE.

- Wang, F., Feng, J., Zhao, Y., Zhang, X., Zhang, S., & Han, J. (2019). Joint activity recognition and indoor localization with WiFi fingerprints. *IEEE Access*, 7(1), 80058–80068. <https://doi.org/10.1109/Access.6287639>.
- Wang, G., Abbasi, A., & Liu, H. (2021a). Dynamic phase calibration method for CSI-based indoor positioning. In *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0108–0113), IEEE.
- Wang, G., Abbasi, A., & Liu, H. (2021b). WiFi-based environment adaptive positioning with transferable fingerprint features. In *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0123–0128), IEEE.
- Wang, H., Li, J., Cui, W., Lu, X., Zhang, Z., Sheng, C., & Liu, Q. (2019). Mobile robot indoor positioning system based on k-ELM. *The Journal of Sensors*, 2019(97), 7547648. <https://doi.org/10.1155/2019/7547648>
- Wang, R., Li, Z., Luo, H., Zhao, F., Shao, W., & Wang, Q. (2019). A robust Wi-Fi fingerprint positioning algorithm using stacked denoising autoencoder and multi-layer perceptron. *Remote Sensing*, 11(11), 1293. <https://doi.org/10.3390/rs11111293>
- Wang, X. (2019). WiFi fingerprinting based indoor localization: When CSI tensor meets deep residual sharing learning. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699.
- Wang, X., Gao, L., & Mao, S. (2015). PhaseFi: Phase fingerprinting for indoor localization with a deep learning approach. In *2015 IEEE Global Communications Conference (GLOBECOM)* (pp. 1–6), IEEE.
- Wang, X., Gao, L., & Mao, S. (2016). CSI phase fingerprinting for indoor localization with a deep learning approach. *IEEE Internet of Things Journal*, 3(6), 1113–1123. <https://doi.org/10.1109/JIOT.2016.2558659>
- Wang, X., Gao, L., & Mao, S. (2017). BiLoc: Bi-modal deep learning for indoor localization with commodity 5 GHz WiFi. *IEEE Access*, 5(1), 4209–4220. <https://doi.org/10.1109/ACCESS.2017.2688362>.
- Wang, X., Gao, L., Mao, S., & Pandey, S. (2015). DeepFi: Deep learning for indoor fingerprinting using channel state information. In *2015 IEEE Wireless Communications and Networking Conference (WCNC)* (pp. 1666–1671), IEEE.
- Wang, X., Gao, L., Mao, S., & Pandey, S. (2016). CSI-based fingerprinting for indoor localization: A deep learning approach. *IEEE Transactions on Vehicular Technology*, 66(1), 763–776. <https://doi.org/10.1109/TVT.2016.2545523>
- Wang, X., Wang, X., & Mao, S. (2017a). CiFi: Deep convolutional neural networks for indoor localization with 5 GHz Wi-Fi. In *2017 IEEE International Conference on Communications (ICC)* (pp. 1–6), IEEE.
- Wang, X., Wang, X., & Mao, S. (2017b). ResLoc: Deep residual sharing learning for indoor localization with CSI tensors. In *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)* (pp. 1–6), IEEE.
- Wang, X., Wang, X., & Mao, S. (2018a). Deep convolutional neural networks for indoor localization with CSI images. *IEEE Transactions on Network Science and Engineering*, 7(1), 316–327. <https://doi.org/10.1109/TNSE.6488902>
- Wang, X., Wang, X., & Mao, S. (2018b). RF sensing in the internet of things: A general deep learning framework. *IEEE Communications Magazine*, 56(9), 62–67. <https://doi.org/10.1109/MCOM.2018.1701277>
- Wang, X., Wang, X., Mao, S., Zhang, J., Periaswamy, S. C., & Patton, J. (2020). Indoor radio map construction and localization with deep Gaussian Processes. *IEEE Internet of Things Journal*, 7(11), 11238–11249. <https://doi.org/10.1109/JIoT.6488907>
- Wang, Y., Gao, J., Li, Z., & Zhao, L. (2020). Robust and accurate Wi-Fi fingerprint location recognition method based on deep neural network. *Applied Sciences*, 10(1), 321. <https://doi.org/10.3390/app10010321>
- Wu, B. F., Jen, C. L., & Chang, K. C. (2007). Neural fuzzy based indoor localization by Kalman filtering with propagation channel modeling. In *2007 IEEE International Conference on Systems, Man and Cybernetics* (pp. 812–817), IEEE.
- Wu, G. S., & Tseng, P. H. (2018). A deep neural network-based indoor positioning method using channel state information. In *2018 International Conference on Computing, Networking and Communications (ICNC)* (pp. 290–294), IEEE.

- Wu, P., Imbiriba, T., LaMountain, G., Vilà-Valls, J., & Closas, P. (2019). WiFi fingerprinting and tracking using neural networks. In *Proceedings of the 32nd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2019)* (pp. 2314–2324), Miami, FL, USA.
- Xiao, L., Behboodi, A., & Mathar, R. (2017). A deep learning approach to fingerprinting indoor localization solutions. In *2017 27th International Telecommunication Networks and Applications Conference (ITNAC)* (pp. 1–7), IEEE.
- Xingli, G., Yaning, L., Ruihui, Z. (2018). Indoor positioning technology based on deep neural networks. In *2018 Ubiquitous Positioning, Indoor Navigation and Location-based Services (UPINLBS)* (pp. 1–6), IEEE.
- Xu, C., Jia, Z., Chen, P., & Wang, B. (2016). CSI-based autoencoder classification for Wi-Fi indoor localization. In *2016 Chinese Control and Decision Conference (CCDC)* (pp. 6523–6528), IEEE.
- Xu, Y., & Sun, Y. (2012). Neural network-based accuracy enhancement method for WLAN indoor positioning. In *2012 IEEE Vehicular Technology Conference (VTC Fall)* (pp. 1–5), IEEE.
- Xu, Y., Zhou, M., & Ma, L. (2009). WiFi indoor location determination via ANFIS with PCA methods. In *2009 IEEE International Conference on Network Infrastructure and Digital Content* (pp. 647–651), IEEE.
- Xue, J., Liu, J., Sheng, M., Shi, Y., & Li, J. (2020). A WiFi fingerprint based high-adaptability indoor localization via machine learning. *China Communications*, 17(7), 247–259. <https://doi.org/10.1109/CC.6245522>
- Zhang, G., Wang, P., Chen, H., & Zhang, L. (2019). Wireless indoor localization using convolutional neural network and Gaussian process regression. *Sensors*, 19(11), 2508. <https://doi.org/10.3390/s19112508>
- Zhang, H., Du, H., Ye, Q., & Liu, C. (2019). Utilizing CSI and RSSI to achieve high-precision outdoor positioning: A deep learning approach. In *ICC 2019-2019 IEEE International Conference on Communications (ICC)* (pp. 1–6), IEEE.
- Zhang, L., Chen, Z., Cui, W., Li, B., Chen, C., Cao, Z., & Gao, K. (2020). Wifi-based indoor robot positioning using deep fuzzy forests. *IEEE Internet of Things Journal*, 7(11), 10773–10781. <https://doi.org/10.1109/JIoT.6488907>
- Zhang, M., Jia, J., Chen, J., Deng, Y., Wang, X., & Aghvami, A. H. (2021). Indoor localization fusing WiFi with Smartphone inertial sensors using LSTM networks. *IEEE Internet of Things Journal*, 8(17), 13608–13623. <https://doi.org/10.1109/JIOT.2021.3067515>
- Zhang, T., & Yi, M. (2018). The enhancement of WiFi fingerprint positioning using convolutional neural network. In *Proceedings of the 2018 International Conference on Computer, Communication and Network Technology*, Wuzhen, China, 29-30 June 2018.
- Zhang, W., Liu, K., Zhang, W., Zhang, Y., & Gu, J. (2014). Wi-Fi positioning based on deep learning. In *2014 IEEE International Conference on Information and Automation (ICIA)* (pp. 1176–1179), IEEE.
- Zhang, W., Liu, K., Zhang, W., Zhang, Y., & Gu, J. (2016). Deep neural networks for wireless localization in indoor and outdoor environments. *Neurocomputing*, 194(C), 279–287. <https://doi.org/10.1016/j.neucom.2016.02.055>.
- Zhang, W., Sengupta, R., Fodero, J., & Li, X. (2017). DeepPositioning: Intelligent fusion of pervasive magnetic field and WiFi fingerprinting for smartphone indoor localization via deep learning. In *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 7–13), IEEE.
- Zhang, Z., Lee, M., & Choi, S. (2020). Deep learning-based indoor positioning system using multiple fingerprints. In *2020 International Conference on Information and Communication Technology Convergence (ICTC)* (pp. 491–493), IEEE.
- Zhao, B., Zhu, D., Xi, T., Jia, C., Jiang, S., & Wang, S. (2019). Convolutional neural network and dual-factor enhanced variational Bayes adaptive Kalman filter based indoor localization with Wi-Fi. *Computer Networks*, 162(1), 106864. <https://doi.org/10.1016/j.comnet.2019.106864>.
- Zhong, Y., Yuan, Z., Zhao, S., & Luo, X. (2018). A Wifi positioning method based on stack auto encoder. In *2018 7th International Conference on Digital Home (ICDH)* (pp. 286–293), IEEE.
- Zhou, C., & Gu, Y. (2017). Joint positioning and radio map generation based on stochastic variational Bayesian inference for FWIPS. In *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 1–10), IEEE.

- Zhou, C., & Wieser, A. (2016). Application of backpropagation neural networks to both stages of fingerprinting based WIPS. In *2016 Fourth International Conference on Ubiquitous Positioning, Indoor Navigation and Location based Services (UPINLBS)* (pp. 207–217), IEEE.
- Zhou, R., Hao, M., Lu, X., Tang, M., & Fu, Y. (2018). Device-free localization based on CSI fingerprints and deep neural networks. In *2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)* (pp. 1–9), IEEE.
- Zhou, Z., Yu, J., Yang, Z., & Gong, W. (2020). MobiFi: Fast deep-learning based localization using mobile WiFi. In *Globecom 2020-2020 IEEE Global Communications Conference* (pp. 1–6), IEEE.
- Zhu, C., Xu, L., Liu, X. Y., & Qian, F. (2018). Tensor-generative adversarial network with two-dimensional sparse coding: Application to real-time indoor localization. In *2018 IEEE International Conference on Communications (ICC)* (pp. 1–6), IEEE.
- Zou, J., Guo, X., Li, L., Zhu, S., & Feng, X. (2018). Deep regression model for received signal strength based WiFi localization. In *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)* (pp. 1–4), IEEE.