

Article

MFA-OSELM Algorithm for WiFi-Based Indoor Positioning System

Ahmed Salih AL-Khaleefa ^{1,*} , Mohd Riduan Ahmad ¹, Azmi Awang Md Isa ¹,
Ahmed AL-Saffar ², Mona Riza Mohd Esa ³ and Reza Firsandaya Malik ⁴

¹ Broadband and Networking (BBNET) Research Group, Centre for Telecommunication and Research Innovation (CeTRI), Fakulti Kejuruteraan Elektronik dan Kejuruteraan Komputer (FKEKK), Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, Durian Tunggal 76100, Melaka, Malaysia; riduan@utem.edu.my (M.R.A.); azmiawang@utem.edu.my (A.A.M.I.)

² Faculty of Computer System and Software Engineering, University Malaysia Pahang (UMP), Gambang 26300, Pahang, Malaysia; ahmed_saffar5@siswa.ukm.edu.my

³ Institute of High Voltage and High Current (IVAT), School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia (UTM), Skudai 81310, Johor Bharu, Malaysia; monariza@utm.my

⁴ Faculty of Computer Science, Universitas Sriwijaya (UNSRI), Inderalaya, Sumatera Selatan 30151, Indonesia; rezafm@unsri.ac.id

* Correspondence: ahmed.salih89@siswa.ukm.edu.my; Tel.: +60-11-2122-2077

Received: 27 February 2019; Accepted: 16 April 2019; Published: 18 April 2019



Abstract: Indoor localization is a dynamic and exciting research area. WiFi has exhibited a tremendous capability for internal localization since it is extensively used and easily accessible. Facilitating the use of WiFi for this purpose requires fingerprint formation and the implementation of a learning algorithm with the aim of using the fingerprint to determine locations. The most difficult aspect of techniques based on fingerprints is the effect of dynamic environmental changes on fingerprint authentication. With the aim of dealing with this problem, many experts have adopted transfer-learning methods, even though in WiFi indoor localization the dynamic quality of the change in the fingerprint has some cyclic factors that necessitate the use of previous knowledge in various situations. Thus, this paper presents the maximum feature adaptive online sequential extreme learning machine (MFA-OSELM) technique, which uses previous knowledge to handle the cyclic dynamic factors that are brought about by the issue of mobility, which is present in internal environments. This research extends the earlier study of the feature adaptive online sequential extreme learning machine (FA-OSELM). The results of this research demonstrate that MFA-OSELM is superior to FA-OSELM given its capacity to preserve previous data when a person goes back to locations that he/she had visited earlier. Also, there is always a positive accuracy change when using MFA-OSELM, with the best change achieved being 27% (ranging from eight to 27% and six to 18% for the TampereU and UJIIndoorLoc datasets, respectively), which proves the efficiency of MFA-OSELM in restoring previous knowledge.

Keywords: indoor positioning system; fingerprint; WiFi; transfer learning; machine learning

1. Introduction

In the current mobile technology age, several new and emerging applications are being adopted to improve communication techniques. In most of these applications, a person is always digitally connected, and their devices receive and transmit data via a protected cloud environment or a device in the same room [1]. Essentially, data regarding location take into account the use of different services used by single devices, which is regarded as hugely important in tracking and supervising devices. The Global Positioning System (GPS) is a satellite-based system that is used to estimate the position of a certain device/user. GPS signals that are broadcast via satellites, for instance, are produced with

low power levels, causing weak signal strength indoors. Thus, the building of systems that can afford satisfactory location estimates indoors is important. Such a system is called an indoor positioning system (IPS). Services that depend on IPSs are known as indoor location-based services (ILBSs) [2]. Instances of such services include ILBSs for travelers who need tools or position guidance services to pinpoint significant exotic sites of interest [3]. Another instance can be found in huge shopping areas and malls [4], where customers generally wish to quickly locate a restaurant or shop that has an outstanding user rating. Additionally, in hospitals and clinics, it is necessary to track the locations of wandering patients in real time in order to locate them quickly and give them timely treatment. Also, in distribution depots, staff requires tracking devices in order to detect the location of goods in real time. Moreover, in art displays, visitors may request to look at the most fascinating paintings located within a museum, and so on. As mentioned, it is not guaranteed that the GPS will work correctly in an intricate indoor region, since satellite signals are often obstructed or very weak. Additionally, in comparison with outdoor areas, greater location resolutions are needed to locate users in indoor areas [2]. This presents new issues in the construction of systems for indoor positioning on the basis of time-crucial constraints, high accuracy requirements, and efficiency of energy use. In fact, new developments in mobile and wireless technology have commenced a new era of indoor positioning. WiFi positioning is presently among the most potent techniques for localization, and is based on smart devices' detection and computing. The reasons for this are the availability of the WiFi infrastructure in nearly every indoor area and the WiFi detection feature available in almost all smart devices.

Among the major issues with WiFi positioning is the way in which the fingerprint model can be adopted every time an update to the environment takes place. An update refers to any change due to the population density in the area, the number of access points (APs) that are active, or signal interference caused by internal and/or external sources. This requires a well-performing model to transfer learning that will facilitate the optimization of previous data and facts in present scenarios [5]. The transfer of knowledge and proficiency among domains is called transfer learning, and improves an associated target task by making use of the data acquired from a source task. It is very often not possible to gather the required data for training from scratch when categorizing tasks, and transfer learning was motivated by this partial training data. By the transfer of weights related to already trained data, this method can assist in considerably lowering the quantity of trainable variables in the target area. Concerning WiFi positioning, several studies [5–8] have used the idea of transfer learning to reduce the costs of training related to changes in the environment. A common environmental change is when there is a change in the quantity of APs in an environment. There can be lasting changes, such as the removal or addition of one AP. In this situation, it is necessary to carry out transfer learning once [9]. A greater change takes place in real life, which consequently implies that transfer learning should be carried out more often. At the same time, there should be a maximum preservation of previous data if needed, so that the previous situation can be replicated later on. This paper concentrates on transfer learning of this sort, and particularly addresses the issue of its efficiency in instances, which include cyclic dynamic factors. Therefore, it is noticed when an individual frequently goes to locations that he/she has visited earlier and where this scenario is viewed as an actual and practical problem. In this condition, the transfer-learning model has to preserve previous data about a previously visited place in order to enable and tag that location as revisited. To the best of the authors' knowledge, this issue has not yet been resolved in the learning transfer literature or in studies of WiFi-based localization. Extreme learning machine (ELM) is viewed as an appropriate learning instrument since it possesses a structure with a hidden layer. This structure is suitable for real-time positioning and needs less training efforts in comparison with other types of neural networks, confirming that ELM can be used as a general approximator [10].

The remainder of this article is structured as follows: Section 2 presents associated works; Section 3 discusses the methodology; Section 4 outlines the experimental work and outcomes; and Section 5 presents the conclusion and future work.

2. Related Works

The previous work reviewed here consists of two types: Work conducted in transfer learning and work on ELM deployment for localization based on WiFi.

Concerning the review of transfer learning, the primary development of localization using manifolds alignment (LuMA) was performed by [11] to handle the issue of transfer learning that happens due to changes in signals based on devices and time. LuMA is categorized as a dimensionality reduction method. This signifies that it possesses the capability to learn and create a map between source and target datasets in a low-dimensional space. Furthermore, this mapping association can be utilized to transfer information from source data to target data. Constraint online sequential extreme learning machine (COSELM), presented by Gu et al. [12], is a new technique for indoor positioning that uses incremental data to revise an old model and handle the issue of fluctuation that frequently occurs when a wireless signal passes over time. This research was assessed for its application to WiFi indoor positioning. Based on this evaluation, it was noted that COSELM had superior performance compared to current batch-learning positioning techniques such as support vector machine (SVM) [13], neural networks (NNs) [14], ELM [15,16], and decision tree (DT) [17]. Nevertheless, this research did not consider the problem of varying the number of observable parameters when the number of APs changes.

Zheng et al. [18] carried out independent research that used the transfer model to manage the environmental variation problems between the running time and data-gathering time of a model [18]. A semi-supervised hidden markov model (HMM) was utilized in this study to transfer a learned prototype from one-time duration to another. This configurable model is called the transferred HMM (TrHMM), and its purpose is to transmit information from the earlier model in order to reduce the calibration effort in the current model. However, this model was not successful in considering the recurring shifts that occurred in situations that required the preservation of old information. Other researchers measured nearby spots to re-compute the new fingerprint [19]. Their prototype sought several previously required received signal strength (RSS) fingerprints and recent feedback within the cut-off region. Furthermore, they employed plane-interpolation to compute the new RSS fingerprint for a particular spot. Nonetheless, whenever a dynamic change occurs, this model requires the measurement of fresh spots to be performed. Likewise, reference [20] carried out research that necessitated the use of present signals in order to measure RSS from other signals used as a reference. This technique needs more pivot points when dealing with environmental changes.

Other researchers chose the online sequential extreme learning machine (OSELM) for WiFi positioning [21]. This is due to the fact that OSELM has a rapid learning pace, which can reduce the time and manpower expenses involved in offline location surveys. More essentially, it has a continuous online learning capability that enables the projected localization technique to adjust to environmental changes in a well-timed and automatic way. Moreover, other researchers have used a combination of weighted extreme learning machine (WELM) and the signal tendency index (STI) for WiFi based positioning that has a normalized fingerprint as its basis. Reference [22] presented two kinds of robust extreme learning machines (RELMs), namely close-to-mean constraint and small-residual constraint, in order to resolve the issues of noisy measurement in IPSs. It is based on whether a precise feature mapping in ELM exists. The second-order cone programming was further applied to provide arbitrary hidden nodes and kernelized formulations of RELM, in that order. This has already been applied in indoor positioning that uses WiFi and offered greater accuracy compared to the basic ELM technique.

Besides the use of ELM in WiFi-based positioning, many experts, such as the authors of [23], have also employed it in positioning based on personal dead reckoning (PDR). Initially, these authors created the PDR positioning procedure as an estimation function. Furthermore, they devised a method based on sliding-window in order to pre-process the acquired inertial sensor information and produce a feature dataset. Lastly, they presented the PDR technique based on the online sequential extreme learning machine (OSELM) to deal with the problem of pedestrian localization. Due to the OSELM

extreme pace of learning and complete ability for location estimation, this technique has the capacity to dynamically adapt to the environment of localization and reduce the positioning errors.

For both the portions, transfer learning and ELM deployment, the research by [9] used the ELM method for the transfer-learning model. The created model can add or remove APs from the positioning environment, which in turn causes modifications to the fingerprint framework. Transfer learning was employed for the NN to adjust to a novel circumstance without the need to accumulate the new fingerprint again. The previous knowledge acquired within the NN can be shifted to the present network with the help of two matrices: One for input-weight transfer and another for the input-weight supplement vector. The second matrix helps the model to carry out the necessary adaptations regarding the evolving aspect of feature matrices amongst the areas including continuous online learning. Moreover, this framework is effective for preventing conventional and arduous training procedures at the time of possible update in the distribution of data due to environmental or domain-related changes. However, it loses all the previous data and knowledge obtained from the network. This data is crucial in situations in which there is a substantial dynamic environmental change, which in turn causes the restoration of the previous data when another change happens in the system. One particular instance concerning WiFi positioning is when the users navigate frequently between the same areas. The work conducted in [24,25] describes a novel type of extreme learning machine using external memory and transfer learning: ITM-OSELM. In these studies, the authors applied ITM-OSELM to WiFi localization, and showed a good improvement in performance in the context of the cyclic dynamic and feature adaptability of WiFi navigation. However, the approach used external memory, which means it could not preserve the knowledge in its neural network without the use of external memory. In [26], the authors transformed the RSS to a standardized location fingerprint based on the Procrustes analysis, and introduced a similar metric, the STI, for matching standardized fingerprints. The article focused on dealing with device heterogeneity and environmental change, and WELM was used with the STI to increase the robustness. However, the approach did not tackle the aspect of knowledge transfer, nor that of preserving previously gained knowledge when it is needed in certain situations, e.g., when visiting a previously visited area.

To summarize, the most widely used methodology determined from the literature review with the greatest potential application as a technique for tackling dynamic environmental changes is the deployment of supplementary measurements or points and the re-estimation of WiFi fingerprints. Moreover, transfer learning has been employed together with many other machine learning approaches. However, it did not indicate any adjustment to the changes in the number of attributes, nor did it preserve previous knowledge, which is the emphasis of this work.

Review of Feature Adaptive Online Sequential Extreme Learning Machine (FA-OSELM)

FA-OSELM can be defined as a method of transferring previous knowledge from a pre-trained neural network to a new network, based on the difference in the number of features in both. This approach was developed by [9]. This model put forward aims to transfer weight values from the old to the new classifier. Assuming that hidden nodes (L) have similar amounts, FA-OSELM provides an input-weight supplement vector Q_i as well as an input-weight transfer matrix P , which allow moving from the old weights a_i to the new weights a'_i with regards to the equation that accounts for the change in the amount of features from m_t to m_{t+1} :

$$\{a'_i = a_i \cdot P + Q_i\}_{i=1}^L \quad (1)$$

where

$$P = \begin{bmatrix} P_{11} & \cdots & P_{1m_{t+1}} \\ \vdots & \ddots & \vdots \\ P_{m_t1} & \cdots & P_{m_t m_{t+1}} \end{bmatrix}_{m_t \times m_{t+1}} \quad (2)$$

$$Q_i = \begin{bmatrix} Q_1 & \dots & Q_{m_{t+1}} \end{bmatrix} 1 \times m_{t+1} \quad (3)$$

where matrix P must adhere to the following rules:

- a. For every line, there is only one '1'; the rest of the values are all '0';
- b. Every column has at most one '1'; the rest of the values are all '0';
- c. $P_{ij} = 1$ signifies that following a change in the feature dimension, the i th dimension of the original feature vector will become the j th dimension of the new feature vector.

When the feature dimension increases, Q_i will function as the supplement. It also adds the corresponding input-weight for the newly added attributes. Furthermore, the rules below apply to Q_i :

- a. Lower feature dimensions indicate that Q_i can be termed as an all-zero vector. Hence, no additional corresponding input weight is required by the newly added features;
- b. In cases where the feature dimension increases, if the new feature is embodied by the i th item of a'_i , a random generation of the i th item of Q_i should be carried out based on the distribution of a_i .

3. Methodology

This section provides a description of the approaches and methods that were used to build the knowledge-preserving model pertaining to WiFi localization. It starts by describing the creation of dynamic scenarios pertaining to localization. Then, the design of the algorithmic and mathematical procedure for the maximum feature adaptive online sequential extreme learning machine (MFA-OSELM) is described. Finally, the evaluation measures are conducted and presented.

3.1. Generating Dynamic Scenarios of Localization

This section describes the creation of dynamic states with regards to a person moving from one place to another in an area, which results in changes in the number of available APs. The aspect of knowledge preservation is crucial when handling mobility scenarios. For instance, a person might move from area A, in which N_1 features (APs) are the only ones available, to area B, in which N_2 features are available. The subsequent return of the person to area A would necessitate the use of previous knowledge. Through FA-OSELM, knowledge is transferred from area A to area B, however there is a chance of knowledge loss due to a change in dimension pertaining to the neural network when the person moves to an area with fewer APs. Next, upon the return of the person to the previous area, the new transferred knowledge neural network cannot return the lost knowledge. As a result, an extended period is needed to gain back the lost previous knowledge. This can be addressed by developing the idea of transferring knowledge to a new model, which can also assure knowledge preservation along with minimal knowledge loss. In the present study, the following scenarios are investigated in order to quantify this limitation of FA-OSELM:

1. The movement of a person from area A to area B. The number of APs in A (N_A) is higher than the number of APs in B (N_B). The APs in B are contained in the APs in A, or $AP_B \subset AP_A$. As shown in Figure 1a, the person then returns to area A. The red line signifies their trajectory from area A to area B, while the blue line denotes their trajectory from area B to area A. Based on the figure, it can be seen that all APs are contained in area A, which explains why $AP_B \subset AP_A$.
2. The person moves from area A to area B. N_A is greater than or equal to N_B . The APs in area B are not contained in the APs in area A, or $AP_A \not\subset AP_B$. Then, the person returns to area A, as represented in Figure 1b. As in the previous scenario, the red line represents the trajectory from area A to area B, while the blue line represents the trajectory from area B to area A. It can be seen that the APs are distributed in both areas A and B, with greater AP density in area A, i.e., $N_A > N_B$ and $AP_A \not\subset AP_B$.
3. The person moves from area A to area B. N_A is less than N_B . The APs in area B are contained in the APs in area A, or $AP_A \subset AP_B$. Then, the person returns to area A, as represented in Figure 1c.

As in scenario 1, the APs are distributed in one area only; however, in this scenario, it is area B. Thus, $AP_A \subset AP_B$.

4. The person moves from area A to area B. N_A is less than or equal to N_B . The APs in area B are not contained in the APs in area A, or $AP_A \not\subset AP_B$. Then, the person returns to area A, as represented in Figure 1d. As in scenario 2, the APs are distributed in both areas, however this time there is a higher AP density in area B.

The four scenarios are summarized in Table 1. N_A denotes number of access points in area A, while N_B denotes number of access points in area B, and AP_A denotes the access points of area A, while AP_B denotes the access points of area B.

Table 1. Dynamic scenarios of localization.

| Scenario | Movement Direction | Number of APs | AP Subset Relation |
|----------|--------------------|----------------|-------------------------|
| S1 | A→B→A | $N_B < N_A$ | $AP_B \subset AP_A$ |
| S2 | A→B→A | $N_B \leq N_A$ | $AP_B \not\subset AP_A$ |
| S3 | A→B→A | $N_A < N_B$ | $AP_A \subset AP_B$ |
| S4 | A→B→A | $N_A \leq N_B$ | $AP_A \not\subset AP_B$ |

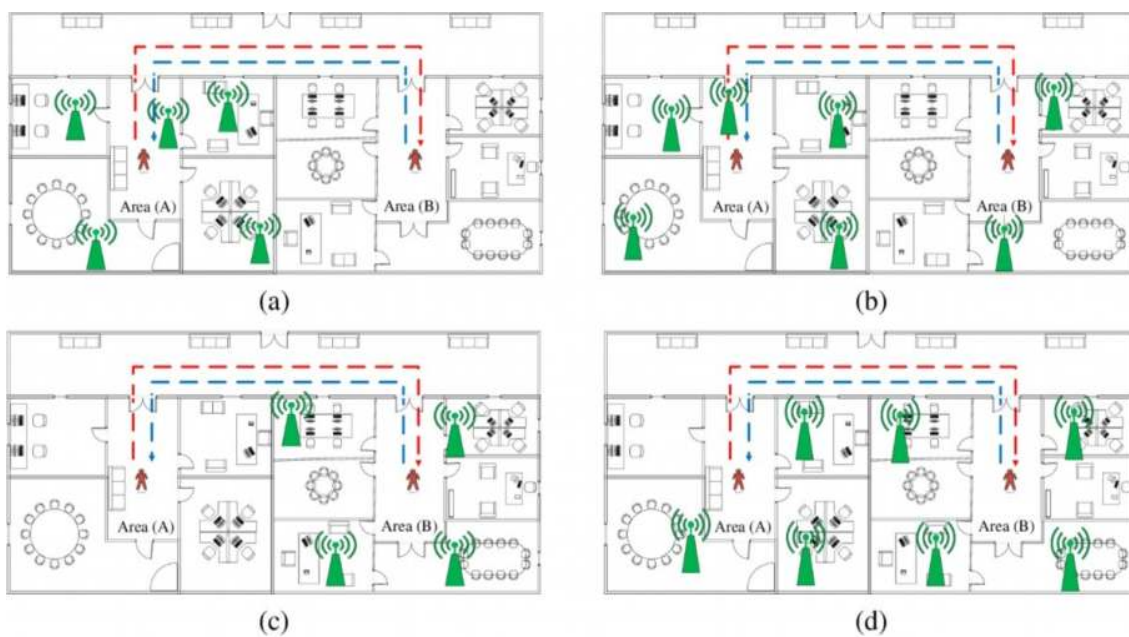


Figure 1. Evaluation scenarios of different areas in an environment according to access point (AP) distribution. (a) Scenario 1: $AP_B \subset AP_A$ and $N_B \leq N_A$; APs are located in area A only; (b) scenario 2: $AP_B \not\subset AP_A$ and $N_B \leq N_A$; APs are located in both areas, with a higher AP density in area A; (c) scenario 3: $AP_A \subset AP_B$ and $N_B \leq N_A$; APs are located in area B only; (d) scenario 4: $AP_A \not\subset AP_B$ and $N_A < N_B$; APs are located in both areas, with a higher AP density in area B.

3.2. General Algorithmic Procedure

The generic process of MFA-OSELM is the same as that of FA-OSELM. However, in order to alter the model to permit knowledge preservation, the transferring matrices and the variant of the used activation function were altered. This process is shown in Figure 2.

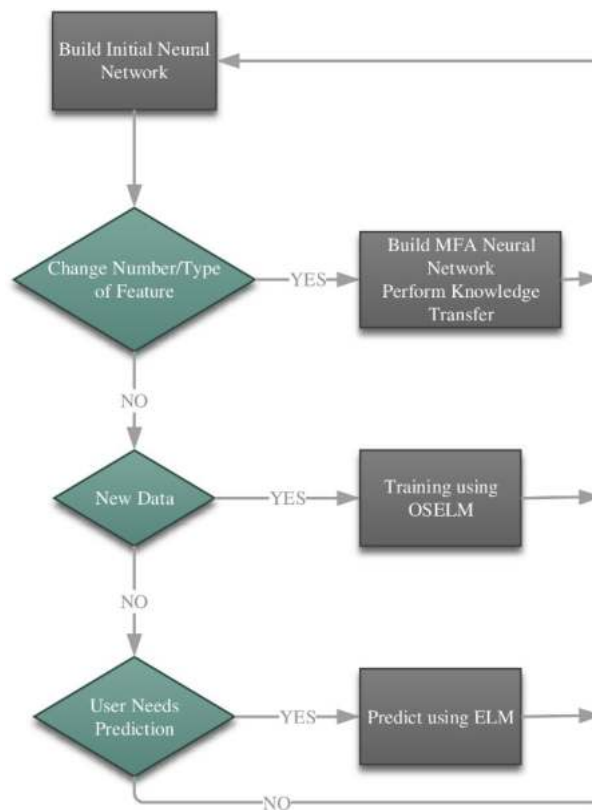


Figure 2. General procedure for the maximum feature adaptive online sequential extreme learning machine (MFA-OSELM) algorithm.

The process of MFA-OSELM allows the constant updating of the neural network whenever data is obtained. Furthermore, it looks for any alteration in the number of features and reacts to such a change by carrying out transfer learning. There are four key operations of MFA-OSELM:

1. Constructing the initial neural network;
2. Constructing the maximum feature adaptive NN and carrying out transfer learning whenever a change in the number of features takes place;
3. Impart training whenever data is accessible;
4. Carry out prediction whenever a user request is required (or intermittently).

In the process of creating the preliminary neural network, the model criteria are chosen from the primary data file of the characteristic d_0 specimen. These criteria are the quantity of concealed nodes L and the activation relation $g(x)$ which fulfills the criteria of passing through $(0, 0)$ as in a tansig, where:

$$\text{tansig}(x) = 2/(1 + e^{-2x}) - 1 \tag{4}$$

It is important to note that the difference between MFA-OSELM and FA-OSLEM is in the approach of combining the input-weight transfer matrix P and the input-weight supplement vector Q_i , $i = 1, 2, \dots, L$.

In FA-OSLEM, P and Q_i are given by the following equations:

$$P = (a_{i,j})_{m_{t-1} \times m_t} \text{ where } \begin{cases} a_{i,j} = 1 \text{ when the feature } i \text{ remains in the new data} \\ a_{i,j} = 0 \text{ when the feature } i \text{ is disabled in the new data} \end{cases} \tag{5}$$

$$Q_i = [Q_j]_{1 \times m_t} \begin{cases} Q_j = 1 \text{ when the feature appears only in the new data} \\ Q_j = 0 \text{ when the feature is only in the old data} \end{cases} \tag{6}$$

where $j = 1, \dots, m_t$.

In MFA-OSELM, P and Q_i are given as follows:

$$P = (a_{i,j})_{m_{t-1} \times M} \begin{cases} a_{i,j} = 1 \text{ when the feature } i \text{ remains in the new data} \\ a_{i,j} = 0 \text{ when the feature } i \text{ is disabled in the new data} \end{cases} \quad (7)$$

$$Q_i = [Q_j]_{1 \times M} \begin{cases} Q_j = 1 \text{ when the feature appears only in the new data} \\ Q_j = 0 \text{ when the feature is only in the old data} \end{cases} \quad (8)$$

where $j = 1, \dots, M$, where:

$$M = \max(m_{t-1}, m_t) \quad (9)$$

As soon as Q_i and P are merged, the learning is transferred in the same way as in FA-OSELM [9]. The learning transfer happens only after the quantity of APs change. If the quantity of APs does not change, then two things can happen: Either there is a user call for prediction where the classical ELM prediction functions are employed or new labeled data become available where the classical OSELM training is applied. Algorithm 1 depicts the pseudo-code for this overall process.

Algorithm 1. Procedure of the maximum feature adaptive online sequential extreme learning machine (MFA-OSELM) for WiFi localization.

- 1- **Input:** Data of access points RSS features $d_t = (x_t^1, x_t^2, \dots, x_t^{m_t}, y_t)$, where x_t^i denotes the RSS signal of access points i , y_t denotes the location of the user at time t , and m_t denotes the number of access points
 - 2- **Output:**
 - 3- Trained neural network based on data of features of access points
 - 4- Predicted location y_t^* when user request for prediction is needed
 - 5- **Start:**
 - 6- For each time t , receive data block $d_t = (x_t^1, x_t^2, \dots, x_t^{m_t}, y_t)$
 - 7- Build single hidden-layer neural network NN_t with size $(m_t \times L, L \times 1)$, where L denotes the number of neurons in the hidden layer
 - 8- If $m_t \neq m_{t-1}$ or $type(x_t^1, x_t^2, \dots, x_t^{m_t}) \neq type(x_{t-1}^1, x_{t-1}^2, \dots, x_{t-1}^{m_{t-1}})$
 - 9- $NN_t = transfer_MFA(NN_{t-1}, m_t, m_{t-1})$
 - 10- Else if $y_t \neq []$
 - 11- $NN_t = train_OSELM(NN_t, d_t)$
 - 12- Else
 - 13- $NN_t = predict_ELM(NN_t, d_t)$
 - 14- End if
 - 15- **End**
-

3.3. Evaluation Measures

This section discusses the standard calculation methods employed for this study. It also discusses how the evolved prototypes were evaluated. The localization system is based on machine learning that performs the classification for any of the possible locations using the fingerprint dataset. The categorization is based on common assessment metrics. For instance, if someone is in location 1 in area A, the classifier classifies it as 1; this is considered to be a true positive (TP). When the classifier classifies it as 2, it is considered to be a false negative (FN). When someone is in location 2 and the classifier classifies it as 1, it is considered to be a false positive (FP). When someone is in location 2 and the classifier classifies it to not be in location 1, it is considered to be a true negative (TN). Out of all of these, i.e., TP, TN, FP, and FN, we can construct a binary confusion matrix, which can be used to compute other performance criteria. These criteria are shown using a statistical matrix known as the confusion matrix. It should be noted that positive and negative classifications are ascertained based on

a specific place. Thus, another type of binary confusion matrix can be formed considering the positive and negative classifications [27].

Moreover, assessment criteria that are able to compute a more detailed difference in performance were added. First, the most accurate localization at a specific place was computed by deriving the maximum temporal accuracy of the localization when someone is in the same location. This measure indicates the relative improvement of one new method over an old one. Basically, the maximum difference in the accuracy between the two approaches is used to describe how much improvement was introduced by the new over the old.

$$MAC = accuracy_{new} - accuracy_{old} \quad (10)$$

where $accuracy_{new}$ denotes the accuracies obtained by the new method, and $accuracy_{old}$ denotes the accuracies obtained by the old method.

Second, in the case of maximum accuracy change (MAC), the difference between the two outcomes is always positive. Here, we refer to the outcomes that specify the maximum deviation of accuracies from each other. This indicated the superiority of the MFA-OSELM in all instances. The resulting figure in results section validates this assumption. The final approach uses the standard deviation (SD) of the sequence of precisions. The method that results in a lower SD is the one that gives more stable learning. This measure is devised in such a way that it is minimized when someone visits a previously visited location, which is tantamount to using previously acquired knowledge in place of gaining it again progressively.

4. Experimental Results

This section discusses examples that validate MFA-OSELM, and compares its performance against that of FA-OSELM. First, the datasets are presented, and then the datasets are pre-processed for further validation. Then, the accuracy of the developed model is obtained and juxtaposed using the baseline method.

4.1. Dataset Description and Pre-Processing

This experiment used the UJIIndoorLoc and TampereU datasets. The UJIIndoorLoc dataset was used as the basis for this experiment. This dataset contains data from three buildings of Jaume I University, Valencia, Spain, each with at least four levels and with areas of about 110,000 m² [28]. This can be utilized for classification purposes, e.g., regression or actual floor and building identification, and e.g., actual estimation of longitude and latitude. The dataset was developed in 2013, with more than 20 distinct users accessing it with 25 Android mobile devices. It contains 19,937 training or reference records and 1111 validation or test records. The dataset also contains an aggregate of 529 attributes with a WiFi fingerprint, which also takes into account the coordinates where the information was obtained and other related information.

The Tampere University (TampereU) dataset is an indoor localization dataset used for testing indoor positioning systems that are dependent on the WLAN/WiFi fingerprint. It was created by E.S. Lohan and J. Talvitie, and has been utilized to test indoor localization techniques [29]. The dataset contains data from two buildings of the Tampere University of Technology, Finland, that have three and four levels. The dataset contains 489 test attributes and 1478 training/reference records for the first building and 312 attributes for the second building. The coordinates (latitude, longitude, and height) and WiFi fingerprints (309 WAPs) are also contained in the dataset.

Tables 2 and 3 present the parameters for both of the datasets (UJIIndoorLoc and TampereU) that were used for the creation of the scenarios. It is evident that a different number of characteristics were selected depending on the scenario and the size of the dataset. Block size indicates the number of measurements prepared for one location query provided to the system; each block contains the records

of various classes. We used 100 and 300 records for TampereU and UJIIndoorLoc respectively which was found to be suitable according to the size of the whole dataset and its rate of classes change.

Table 2. Settings of testing scenarios for block size = 300 and regularization parameter $C = 2^{-9}$ for the UJIIndoorLoc dataset.

| Parameter Name | Parameter Description | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 |
|----------------|---|------------|------------|------------|------------|
| N_A | Number of features in location A | 300 | 300 | 150 | 150 |
| N_B | Number of features in location B | 150 | 150 | 300 | 300 |
| $NCF-AB$ | Number of common features between locations A and B | 150 | 99 | 150 | 80 |
| L | Number of hidden neurons | 850 | 850 | 850 | 850 |

Table 3. Settings of testing scenarios for block size = 100 and regularization parameter $C = 2^{-6}$ for the TampereU dataset.

| Parameter Name | Parameter Description | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 |
|----------------|---|------------|------------|------------|------------|
| N_A | Number of features in location A | 100 | 100 | 50 | 50 |
| N_B | Number of features in location B | 50 | 50 | 100 | 100 |
| $NCF-AB$ | Number of Common features between locations A and B | 50 | 11 | 50 | 15 |
| L | Number of hidden neurons | 750 | 750 | 750 | 750 |

For the selection of the scenario parameters, the regularization parameter was dependent on the relationship between the accuracy and regularization parameter (C), and was selected based on $C = 2^{-9}$ for the UJIIndoorLoc dataset. As depicted in Figure 3, the number of hidden neurons L was chosen to be 850 for this dataset, on the basis of the relationship between L and the accuracy. Similarly, L was selected to be 750 and $C = 2^{-6}$ for the TampereU dataset, as shown in Figure 4.

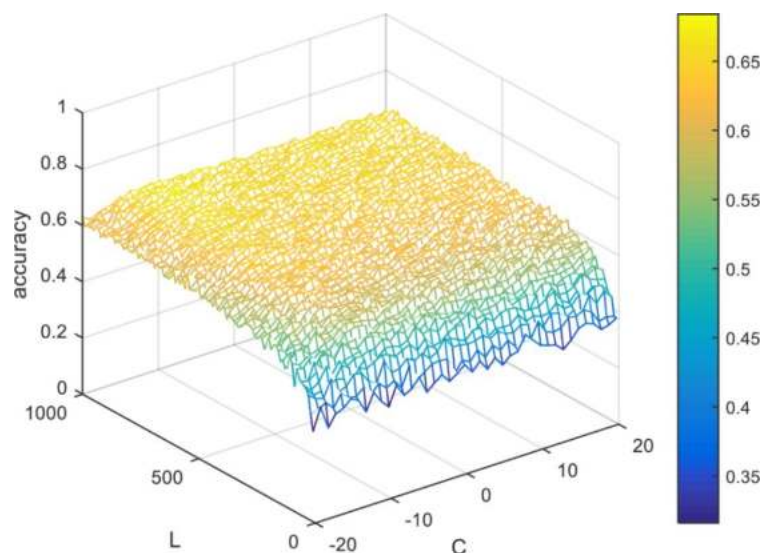


Figure 3. Mesh representation of the relationship between the accuracy, the number of neurons L , and the regularization parameter C for the extreme learning machine (ELM) classifier based on the UJIIndoorLoc dataset.

During the formation of the dataset, the first classification was made with two equal parts, one for testing and another for training. Subsequently, for location A, a set of arbitrary characteristics was selected. Out of this set of characteristics, another set of arbitrary characteristics was selected for

location B. Moreover, to disable the non-working APs, the values of the pre-processing step were set in the range of 0 to 100. Algorithm 2 explains the specifics of the process used for testing scenario 1. The process was applied in the same way for the other scenarios.

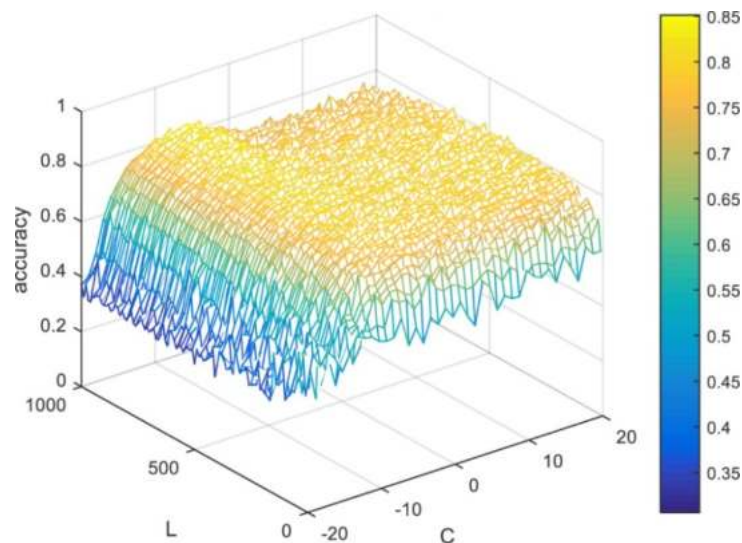


Figure 4. Mesh representation of the relationship between the accuracy, the number of neurons L , and the regularization parameter C for the extreme learning machine (ELM) classifier based on the TampereU dataset.

Algorithm 2. Pseudocode for the dataset preparation for testing scenario 1.

```

1- Input:
2- trainD: Raw training dataset of Wi-Fi fingerprint
3- testD: Raw testing dataset of Wi-Fi fingerprint
4- A, B: two locations with common features
5- NA: number of features in location A
6- NB: number of features in location B
7- Output:
8- DA: data of location A
9- DB: data of location B
10- Start:
11- Replace all 100 in trainD and testD with 0 // preprocessing
12- nData = number of records in trainD
13- number of RecordsA = nData/2
14- testDataA = testD
15- testDataB = testD //testData is the same for locations A and B
16- nFeatures = number of features in dataset
17- nFeatures = 1:nFeatures //create vector of features labels 1, 2,..., nFeatures
18- featuresA = getRandom(features,NA) //select NA random features from features
19- featuresB = getRandom(featuresA,NB) //select NB random features from featuresA
20- trainDataA = trainData(1:numberofRecordsA)
21- trainDataB = trainData(numberofRecordsA + 1:end)
22- trainDataAf = GenerateData(trainDataA, featuresA)
23- testDataAf = GenerateData(testDataA, featuresA)
24- trainDataBf = GenerateData(trainDataB, featuresB)
25- testDataBf = GenerateData(testDataB, featuresB)
26- end

```

4.2. Results and Analysis

To perform the analysis, the measures were generated for all the presented scenarios. Section 4.2.1 presents the details related to the accuracy of the models. The objective of this is to distinguish the accuracy of the developed model from those of another benchmark model with respect to the performance of location prediction. A new measure, MAC, is introduced in Section 4.2.2 to provide a quantitative summary of the differences in accuracy. The standard deviation measure included in Section 4.2.3 was used to analyze the stability of the classifier. The other measures used for the classification evaluation are discussed in Section 4.2.4.

4.2.1. Accuracy

The accuracy of the MFA-OSELM model was computed and compared with that of the FA-OSELM model. The UJIIndoorLoc and TampereU datasets were used to perform the comparison. Accuracy indicates the complementarity of the error ratio to 1, and the chunk indicates a block of data obtained by the classifier for estimation, where the correct values are used for the training of the next chunk (block). Figures 5–8 represent the accuracy in relation to the chunks for all scenarios and both datasets.

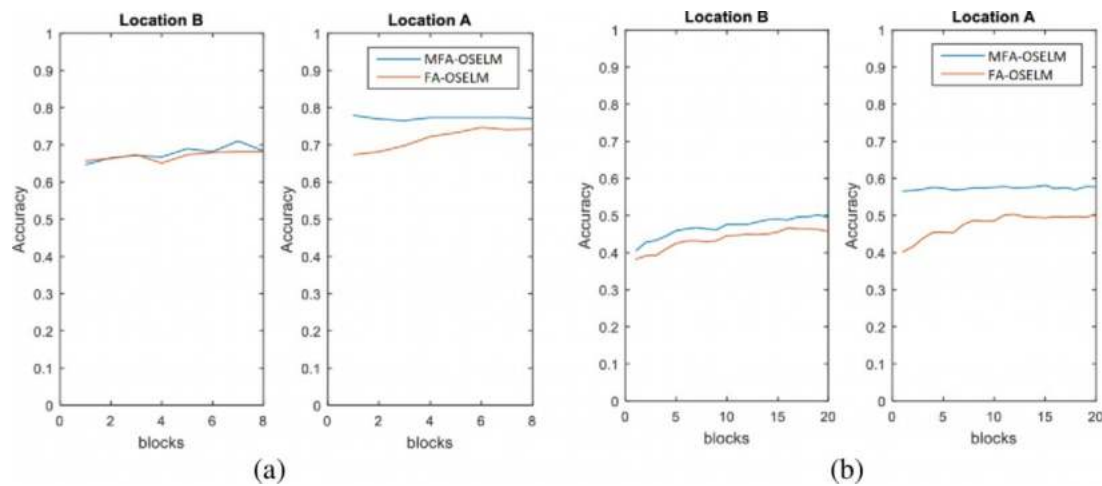


Figure 5. Scenario 1: The accuracy of the maximum feature adaptive online sequential extreme learning machine (MFA-OSELM) and the feature adaptive online sequential extreme learning machine (FA-OSELM) for (a) the TampereU dataset and (b) the UJIIndoorLoc dataset.

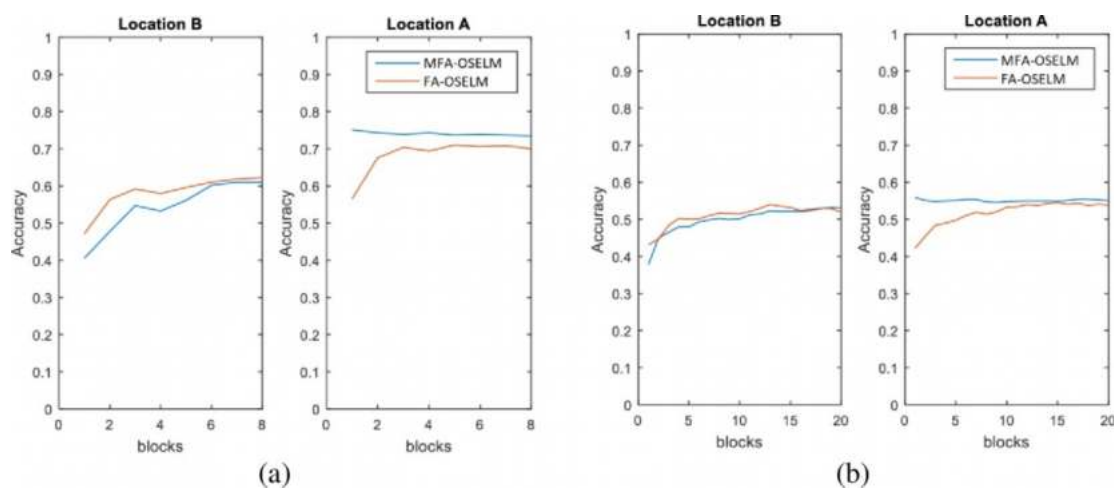


Figure 6. Scenario 2: the accuracy of MFA-OSELM and FA-OSELM for (a) the TampereU dataset and (b) the UJIIndoorLoc dataset.

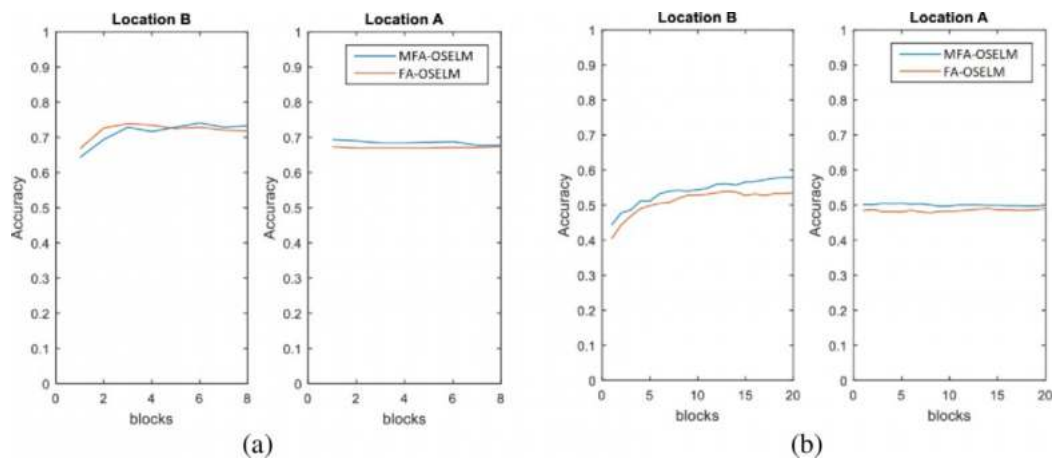


Figure 7. Scenario 3: The accuracy of MFA-OSELM and FA-OSELM for (a) the TampereU dataset and (b) the UJIIndoorLoc dataset.

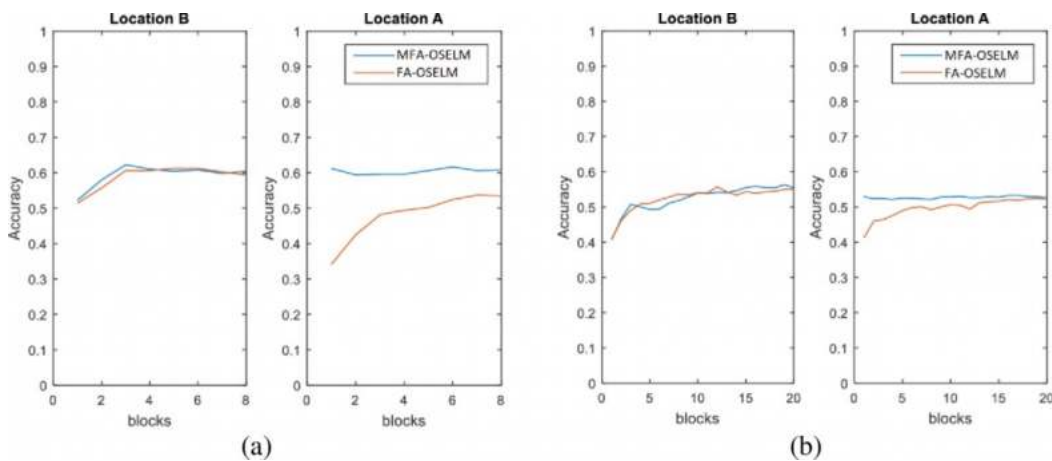


Figure 8. Scenario 4: The accuracy of MFA-OSELM and FA-OSELM for (a) the TampereU dataset and (b) the UJIIndoorLoc dataset.

The graphs in Figures 5–8 show that, for area A, MFA-OSELM and FA-OSELM had similar performances. This is due to the fact that the models did not have previous knowledge to remember. However, in area B, MFA-OSELM was superior to FA-OSELM, which was attributed to the difference in their knowledge preservation aspects. FA-OSELM had the transfer-learning capability. However, since transfer learning is a Markov-type process, it only remembers a previous state and brings its values to the current time. However, MFA-OSELM can restore previous knowledge whenever necessary.

4.2.2. Maximum Accuracy Change

An estimation of MAC was carried out in order to gain a thorough understanding of the variation in the performance of the two models. Figure 9 presents the MAC values for both of the models for scenarios 1 to 4 for both datasets. As can be seen, the values of MAC are positive, which implies that the MFA-OSELM model delivers higher accuracy than the FA-OSELM model. The MAC measure is analyzed in area A, in all scenarios, as it is visited twice, and the purpose of the investigation is to assess the performance of the classification during the second visit. In all scenarios, it is evident that the performance of MFA-OSELM is better than that of FA-OSELM since there is a dearth of knowledge preservation in FA-OSELM, which is not the case with MFA-OSELM. Higher MAC is achieved for scenario 1, with a value of 0.1638 for the UJIIndoorLoc dataset, and for scenario 4, with a value of 0.2714 for the TampereU dataset. This superiority in performance is due to the fact that the MFA-OSELM model has the capability to retain and preserve knowledge. Among all scenarios,

the lowest MAC is achieved for scenario 3 for both datasets. This result is justified by $N_A < N_B$ and $AP_A \subset AP_B$, which shows that FA-OSELM is able to preserve the previous knowledge gained in area A, similar to MFA-OSELM.

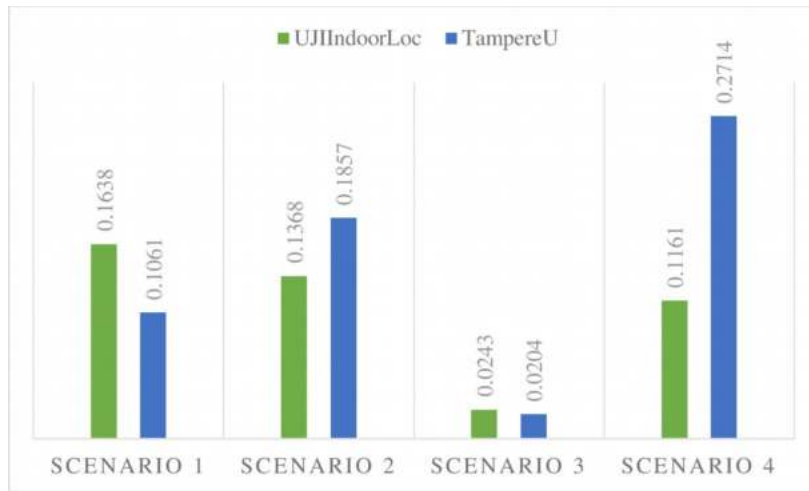


Figure 9. Maximum accuracy change for MFA-OSELM and FA-OSELM for all scenarios for the UJIIndoorLoc and TampereU datasets.

4.2.3. Standard Deviation

The amount of new knowledge acquired by the classifier in an area is indicated by the standard deviation of the model accuracy in relation to the chunks of data. The classifier has the ability to restore previous knowledge when a person goes to an already visited area.

In both the MFA-OSELM and FA-OSELM models, the consistency and stability of the performance was measured using the standard deviations of the accuracy pertaining to the chunks of data, which indicate the volume of new knowledge acquired by the classifier in the area. The standard deviations of signal accuracy for all scenarios are shown in Figures 10 and 11. The computation of these standard deviations was carried out in connection with the learning blocks. Of the two models, FA-OSELM results in the largest standard deviation in all scenarios. This is mainly due to the fact that FA-OSELM does not have the ability to conserve knowledge in accordance with changes in the number of features.

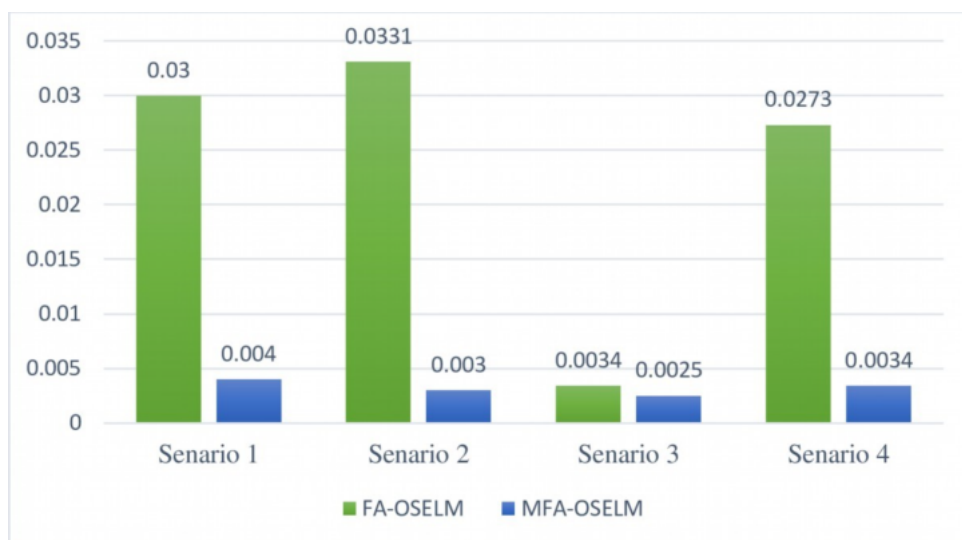


Figure 10. Standard deviations of model signal accuracy obtained using MFA-OSELM and FA-OSELM for the UJIIndoorLoc dataset.

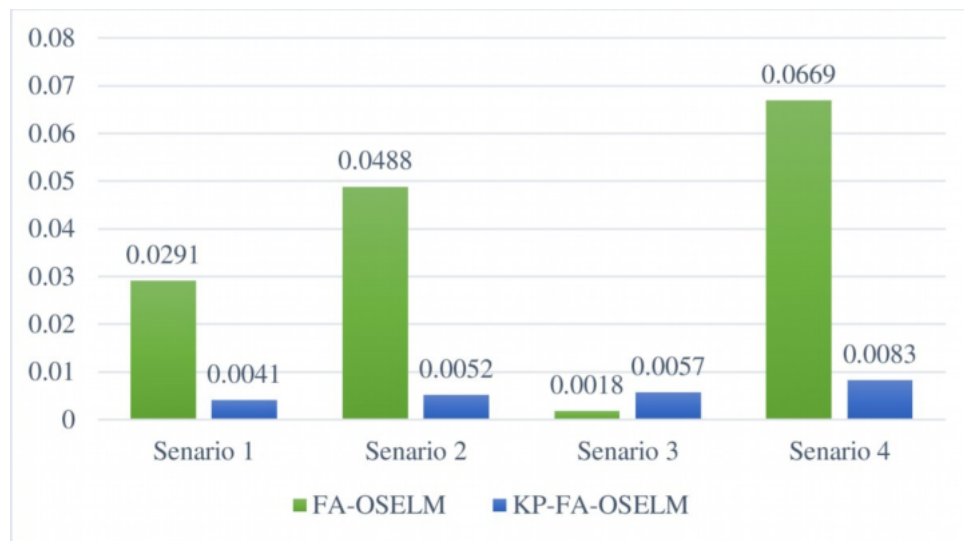


Figure 11. Standard deviations of model signal accuracy obtained using MFA-OSELM and FA-OSELM for the TampereU dataset.

4.2.4. Classification Evaluation Measures

In these measures, the positive and negative identifications of the developed model (i.e., TP, TN, FP, and FN) and their frequencies are presented, and the performance of the benchmark model is evaluated. The recall, precision, and F1 measures for both the MFA-OSELM and FA-OSELM models for all scenarios are presented in Figures 12–15 for the TampereU dataset and in Figures 16–19 for the UJIIndoorLoc dataset. For the UJIIndoorLoc dataset, the MFA-OSELM model generally returns higher frequencies of TP and FP and generally returns lower frequencies of FN, while for the TampereU dataset, the MFA-OSELM model generally returns higher frequencies of TP and TN and generally returns lower frequencies of FP and FN. This suggests that the MFA-OSELM classifier is capable of a superior performance in general. The superiority in performance is observed for all scenarios. Moreover, the F1 and precision parameters are always greater for MFA-OSELM than for FA-OSELM, while the recall parameter is higher for MFA-OSELM in all cases other than for scenario 2 using the UJIIndoorLoc dataset and scenario 4 using the TampereU dataset. With regards to classification, these results confirm that MFA-OSELM is improved relative to FA-OSELM.

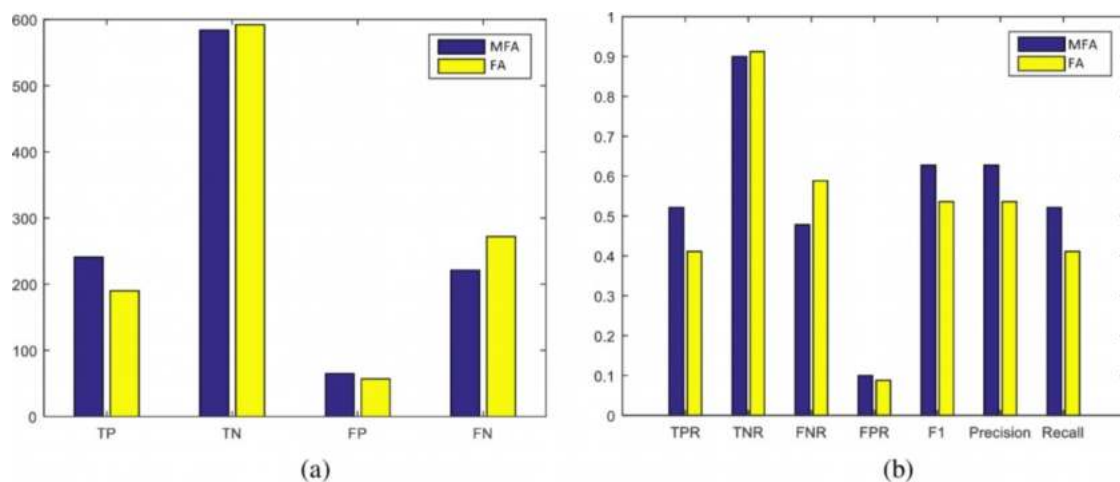


Figure 12. Scenario 1: statistical evaluation of MFA-OSELM and FA-OSELM using two sets of measures, (a,b), for the UJIIndoorLoc dataset.

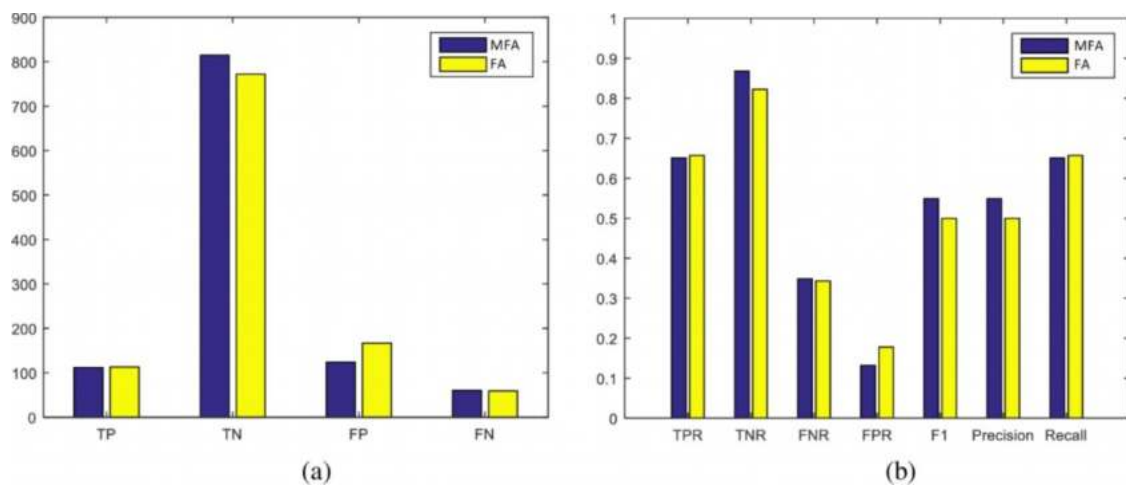


Figure 13. Scenario 2: Statistical evaluation of MFA-OSELM and FA-OSELM using two sets of measures, (a,b), for the UJIIndoorLoc dataset.

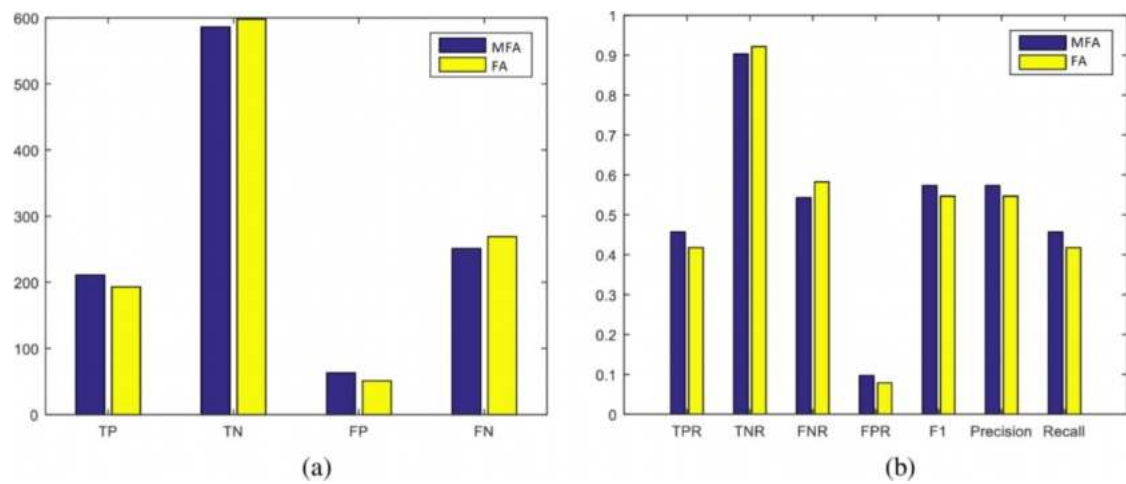


Figure 14. Scenario 3: Statistical evaluation of MFA-OSELM and FA-OSELM using two sets of measures, (a,b), for the UJIIndoorLoc dataset.

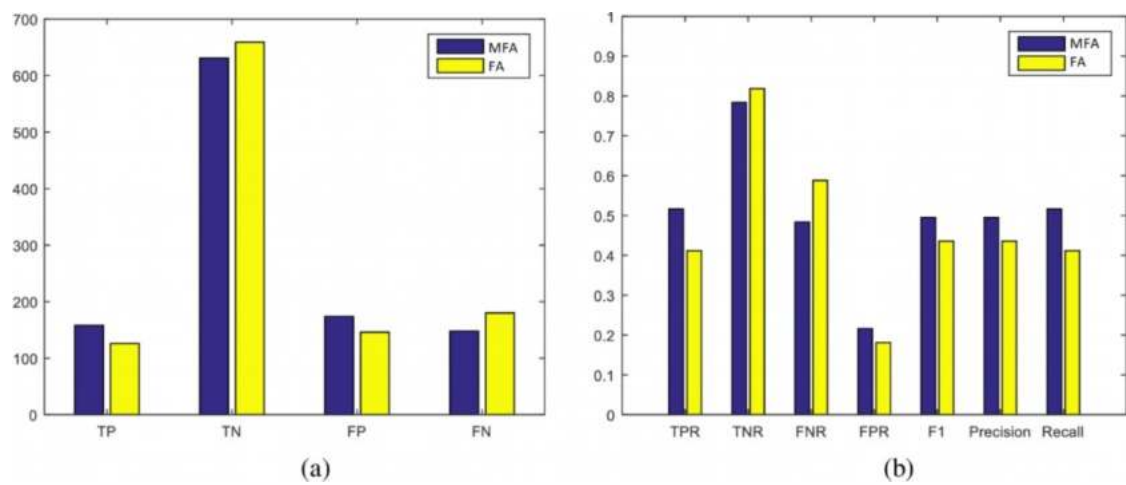


Figure 15. Scenario 4: Statistical evaluation of MFA-OSELM and FA-OSELM using two sets of measures, (a,b), for the UJIIndoorLoc dataset.

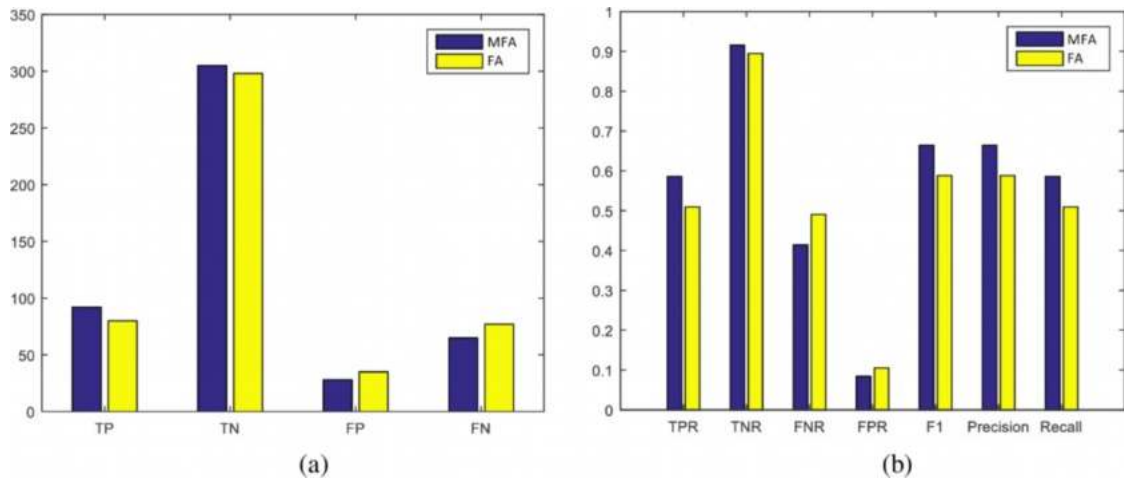


Figure 16. Scenario 1: Statistical evaluation of MFA-OSELM and FA-OSELM using two sets of measures, (a,b), for the TampereU dataset.

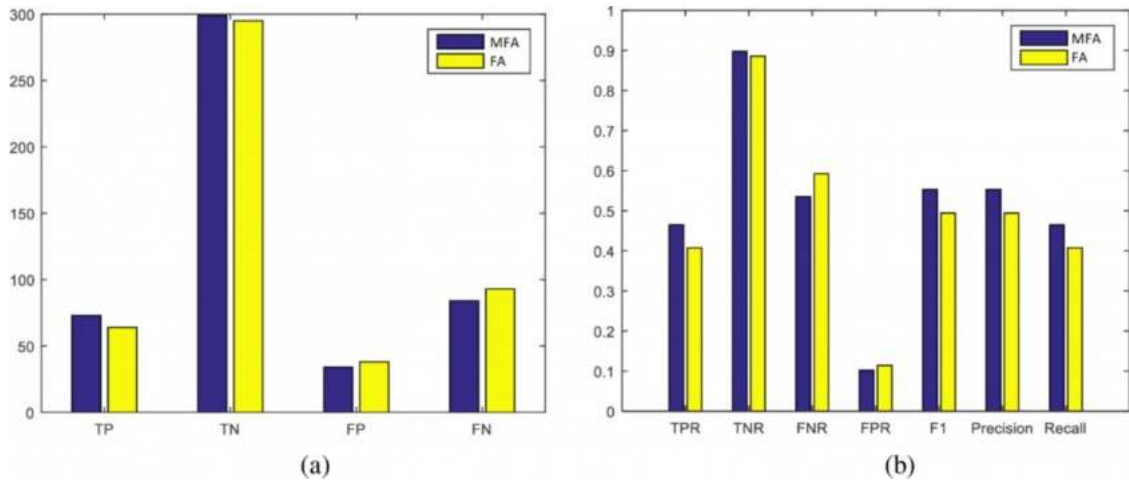


Figure 17. Scenario 2: Statistical evaluation of MFA-OSELM and FA-OSELM using two sets of measures, (a,b), for the TampereU dataset.

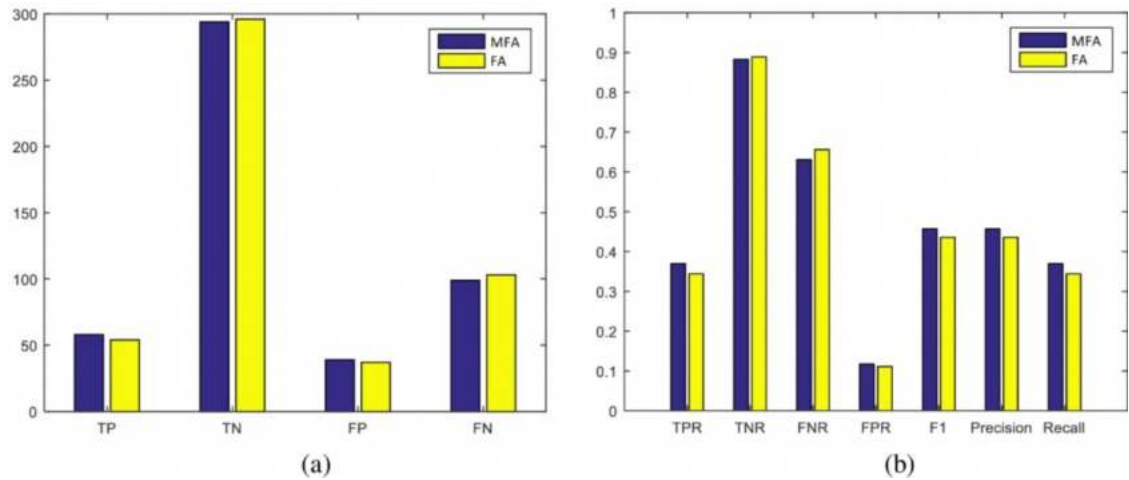


Figure 18. Scenario 3: Statistical evaluation of MFA-OSELM and FA-OSELM using two sets of measures, (a,b), for the TampereU dataset.

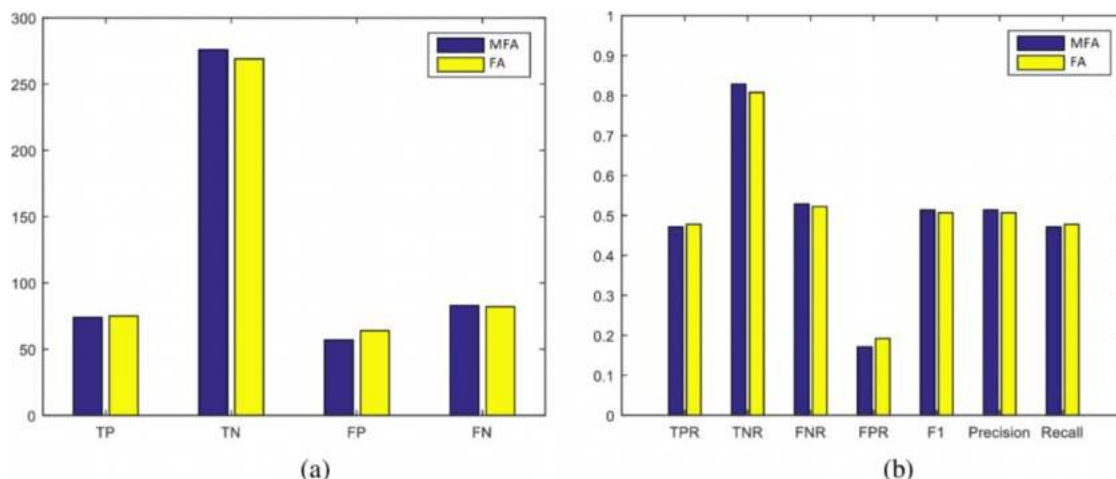


Figure 19. Scenario 4: Statistical evaluation of MFA-OSELM and FA-OSELM using two sets of measures, (a,b), for the TampereU dataset.

5. Summary and Future Work

This article focuses on an online approach for the localization of an individual who is using WiFi technology and moving in an indoor environment. The approach is based on the concept of transfer learning, which has the potential to adapt to a varying number of APs. This study resolves the problem of the cyclic dynamic factor that occurs due to an individual's movements within an indoor environment when the individual iteratively moves to positions that have already been visited. A neural network was required to retain previous knowledge for future use. Based on the analysis carried out in this study, the developed MFA-OSELM model shows a superior performance compared to the FA-OSELM model, with a maximum accuracy increase of 27% observed over the baseline approach (overall, accuracy increases ranged from eight to 27% and from six to 18% for the TampereU and UJIIndoorLoc datasets, respectively). This increase in accuracy convincingly validates the efficiency of the MFA-OSELM model compared to that of the FA-OSELM model. The standard deviation of model accuracy was also used for the performance evaluation, and was almost negligible in all scenarios for the MFA-OSELM model, thereby establishing that this model has an improved stability and is less prone to the loss of saved knowledge compared to the FA-OSELM model.

The objective of a future study will be to assess the influence of common characteristics between the two places on knowledge conservation. The purpose of such study will be to generalize this approach in order to optimize knowledge preservation, which will further involve the cyclic dynamic awareness of other kinds of learners.

Author Contributions: Conceptualization, A.S.A.-K.; Formal analysis, A.S.A.-K.; Funding acquisition, M.R.A., A.A.M.I., M.R.M.E., and R.F.M.; Methodology, A.S.A.-K.; Project administration, A.S.A.-K.; Resources, A.A.-S.; Software, A.S.A.-K.; Supervision, M.R.A., A.A.M.I., and M.R.M.E.; Validation, A.A.-S.; Visualization, A.S.A.-K.; Writing—original draft, A.S.A.-K.; Writing—review and editing, A.S.A.-K., M.R.A., A.A.M.I., A.A.-S., M.R.M.E., and R.F.M.

Funding: This research was funded by the Malaysian Ministry of Education, Universiti Teknikal Malaysia Melaka, under grant PJP/2018/FKEKK(3B)/S01615, and in part by Universiti Teknologi Malaysia under grants 14J64 and 4F966.

Conflicts of Interest: The authors declare no conflict of interest regarding this paper.

References

1. Wan, J.; Zou, C.; Ullah, S.; Lai, C.F.; Zhou, M.; Wang, X. Cloud-enabled wireless body area networks for pervasive healthcare. *IEEE Netw. Mag.* **2013**, *27*, 56–61. [[CrossRef](#)]
2. Xiao, J.; Zhou, Z.; Yi, Y. A Survey on Wireless Indoor Localization from the Device Perspective. *ACM Comput. Surv. CSUR* **2016**, *49*, 25. [[CrossRef](#)]

3. Umanets, A.; Ferreira, A.; Leite, N. GuideMe—A Tourist Guide with a Recommender System and Social Interaction. *Procedia Technol.* **2014**, *17*, 407–414. [[CrossRef](#)]
4. Wang, S.; Fidler, S.; Urtasun, R. Lost shopping! Monocular localization in large indoor spaces. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 2695–2703. [[CrossRef](#)]
5. Zou, H.; Zhou, Y.; Jiang, H.; Huang, B.; Xie, L.; Spanos, C. Adaptive localization in dynamic indoor environments by transfer kernel learning. In Proceedings of the Wireless Communications and Networking Conference (WCNC), San Francisco, CA, USA, 19–22 March 2017; pp. 1–6. [[CrossRef](#)]
6. He, S.; Lin, W.; Chan, S.G. Indoor Localization and Automatic Fingerprint Update with Altered AP Signals. *IEEE Trans. Mob. Comput.* **2017**, *16*, 1897–1910. [[CrossRef](#)]
7. Jiang, X.; Chen, Y.; Liu, J.; Gu, Y.; Chen, Z. Wi-Fi and Motion Sensors Based Indoor Localization Combining ELM and Particle Filter. In *Proceedings of the ELM-2014*; Cao, J., Mao, K., Cambria, E., Man, Z., Toh, K.A., Eds.; Springer: Cham, Switzerland, 2015; Volume 2, pp. 105–113. [[CrossRef](#)]
8. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
9. Jiang, X.; Liu, J.; Chen, Y.; Liu, D.; Gu, Y.; Chen, Z. Feature Adaptive Online Sequential Extreme Learning Machine for lifelong indoor localization. *Neural Comput. Appl.* **2016**, *27*, 215–225. [[CrossRef](#)]
10. Huang, G.; Zhu, Q.; Siew, C. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [[CrossRef](#)]
11. Sun, Z.; Chen, Y.; Qi, J.; Liu, J. Adaptive Localization through Transfer Learning in Indoor Wi-Fi Environment. In Proceedings of the 2008 Seventh International Conference on Machine Learning and Applications Adaptive, San Diego, CA, USA, 11–13 December 2008; pp. 331–336. [[CrossRef](#)]
12. Gu, Y.; Liu, J.; Chen, Y.; Jiang, X. Constraint Online Sequential Extreme Learning Machine for lifelong indoor localization system. In Proceedings of the International Joint Conference on Neural Networks, Beijing, China, 6–11 July 2014; pp. 732–738. [[CrossRef](#)]
13. Mundo, L.B.D.; Ansay, R.L.D.; Festin, C.A.M.; Ocampo, R.M. A Comparison of Wireless Fidelity (Wi-Fi) Fingerprinting Techniques. In Proceedings of the 2011 International Conference on ICT Convergence (ICTC), Seoul, Korea, 28–30 September 2011; pp. 20–25. [[CrossRef](#)]
14. Heidari, M.; Alsindi, N.A.; Pahlavan, K. UDP identification and error mitigation in ToA-based indoor localization systems using neural network architecture. *IEEE Trans. Wirel. Commun.* **2009**, *8*, 3597–3607. [[CrossRef](#)]
15. Huang, G.; Zhu, Q.; Siew, C. Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks. In Proceedings of the 2004 IEEE International Joint Conference on Neural Networks, Budapest, Hungary, 25–29 July 2004; Volume 2, pp. 985–990. [[CrossRef](#)]
16. Huang, G.; Zhou, H.; Ding, X.; Zhang, R. Extreme Learning Machine for Regression and Multiclass Classification. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2012**, *42*, 513–529. [[CrossRef](#)] [[PubMed](#)]
17. Yim, J. Introducing a decision tree-based indoor positioning technique. *Expert Syst. Appl.* **2008**, *34*, 1296–1302. [[CrossRef](#)]
18. Zheng, V.W.; Xiang, E.W.; Yang, Q.; Shen, D. Transferring localization models over time. In Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, Chicago, IL, USA, 13–17 July 2008; pp. 1421–1426.
19. Koweerawong, C.; Wipusitwarakun, K.; Kaemarungsi, K. Indoor localization improvement via adaptive RSS fingerprinting database. In Proceedings of the International Conference on Information Networking 2013, Bangkok, Thailand, 28–30 January 2013; Volume 1, pp. 412–416. [[CrossRef](#)]
20. Lo, C.C.; Hsu, L.Y.; Tseng, Y.C. Adaptive radio maps for pattern-matching localization via inter-beacon co-calibration. *Pervasive Mob. Comput.* **2012**, *8*, 282–291. [[CrossRef](#)]
21. Zou, H.; Jiang, H.; Lu, X.; Xie, L. An Online Sequential Extreme Learning Machine Approach to WiFi Based Indoor Positioning. *IEEE World Forum Internet Things* **2014**, *2014*, 111–116. [[CrossRef](#)]
22. Lu, X.; Zou, H.; Zhou, H.; Xie, L.; Huang, G.B. Robust Extreme Learning Machine With its Application to Indoor Positioning. *IEEE Trans. Cybern.* **2016**, *46*, 194–205. [[CrossRef](#)] [[PubMed](#)]
23. Zhang, M.; Wen, Y. Pedestrian Dead-Reckoning Indoor Localization Based on OS-ELM. *IEEE Access* **2018**, *6*, 6116–6129. [[CrossRef](#)]

24. AL-Khaleefa, A.S.; Ahmad, M.R.; Isa, A.A.M.; Esa, M.R.M.; AL-Saffar, A.; Hassan, M.H. Feature Adaptive and Cyclic Dynamic Learning Based on Infinite Term Memory Extreme Learning Machine. *Appl. Sci.* **2019**, *9*, 895. [[CrossRef](#)]
25. Al-Khaleefa, A.S.; Ahmad, M.R.; Isa, A.A.M.; Esa, M.R.M.; Al-Saffar, A.; Aljeroudi, Y. Infinite-Term Memory Classifier for Wi-Fi Localization Based on Dynamic Wi-Fi Simulator. *IEEE Access* **2018**, *6*, 54769–54785. [[CrossRef](#)]
26. Zou, H.; Huang, B.; Lu, X.; Jiang, H.; Xie, L. A Robust Indoor Positioning System Based on the Procrustes Analysis and Weighted Extreme Learning Machine. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 1252–1266. [[CrossRef](#)]
27. Han, J.; Kamber, M.; Pei, J. 8—Classification: Basic Concepts. In *Data Mining: Concepts and Techniques*; Han, J., Kamber, M., Pei, J., Eds.; Elsevier: Boston, MA, USA, 2011; pp. 327–391, ISBN 978-0-12-381479-1.
28. Torres-Sospedra, J.; Montoliu, R.; Martinez-Uso, A.; Avariento, J.P.; Arnau, T.J.; Benedito-Bordonau, M.; Huerta, J. UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems. In Proceedings of the 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Busan, Korea, 27–30 October 2014; pp. 261–270. [[CrossRef](#)]
29. Lohan, S.E.; Torres-Sospedra, J.; Leppäkoski, H.; Richter, P.; Peng, Z.; Huerta, J. Wi-Fi crowdsourced fingerprinting dataset for indoor positioning. *Data* **2017**, *2*, 32. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).