Bulletin of Electrical Engineering and Informatics
Vol. 13, No. 6, December 2024, pp. 4193~4206
ISSN: 2302-9285, DOI: 10.11591/eei.v13i6.7157

A study on the impact of layout change to knowledge distilled indoor positioning systems

Aqilah Mazlan[1], Yin Hoe Ng[1], Chee Keong Tan[2]
[1]Faculty of Engineering, Multimedia University, Cyberjaya, Malaysia
[2]School of Information Technology, Monash University Malaysia, Subang Jaya, Malaysia

Article Info

Article history:

Received Jul 7, 2023
Revised Apr 22, 2024
Accepted May 16, 2024

Keywords:

Convolutional neural network
Fingerprinting
Indoor positioning system
Knowledge distillation
Layout change

ABSTRACT

Convolutional neural networks (CNN)-based indoor positioning systems (IPS) have gained significant attention over the past decade due to their ability to provide precise localization accuracy. However, the use of CNNs in these systems comes with a higher computational cost. To tackle this issue, recent studies have introduced knowledge distilled positioning schemes to mitigate the computational burden. Despite the clear possibility of performance degradation due to signal fluctuations, there remains a lack of investigation into the performance of knowledge distilled and CNN based indoor positioning schemes in dynamic indoor environment. To fill this research gap, this paper investigates the practicality of implementing knowledge distilled-based indoor positioning schemes in real-world by analyzing the impact of indoor layout change on these schemes. Results demonstrate that in the case of layout change, the knowledge distilled-based indoor positioning schemes without teaching assistant can still achieve good performance, with an improvement of 11.56% in average positioning error compared to simple CNN model, while taking only 49.05% of the complex CNN model's execution time. However, the knowledge distilled-based indoor positioning scheme with teaching assistant fails under the same condition as the inclusion of teacher assistant leads to increased error in modeling the received signal strengths (RSS) and locations relationship.

Corresponding Author:

Yin Hoe Ng
Faculty of Engineering, Multimedia University
Cyberjaya 63100, Malaysia
Email: yhng@mmu.edu.my

1. INTRODUCTION

The demand for pragmatic and efficient solutions to everyday problems has risen as people seek for higher living standards. The deployment of numerous context-aware services and protocols is made possible through accurate real-time indoor localization of users and devices [1]. As a result, research efforts on indoor positioning and localization significantly increased over the previous decade [2]. The limitations of the global positioning systems in a complex indoor environment have prompted the exposure of other positioning technologies and techniques, including the fingerprinting approach which is currently favored by many researchers. A fingerprint-based localization system requires a pre-constructed database to predict the location of the user or the device [3] and thus, this technique is realized in two phases which is the offline training phase and the online localization phase. During the offline phase, a radio map of indoor environment is established by collecting fingerprint information at every point of interest known as the reference point (RP) which are evenly distributed throughout the indoor space. At each location, the mobile device will receive packets of data transmitted from every detectable access point (AP) or beacon. In the localization phase, real-time fingerprints collected at the user's current position is compared with the fingerprints in the database using a pattern recognition or matching algorithm to select the most suitable RP as the predicted location. The wireless technology used for positioning comprises radio frequency identification (RFID), Wi-Fi, Bluetooth, Zigbee [4], [5] long range (LoRa) [6], and ultra-wideband (UWB) [7] because the radio waves can easily travel through walls and human bodies [8]. It is also noted that among the previously mentioned radio technologies, fingerprint-based indoor positioning systems (IPS) predominantly utilize Wi-Fi and Bluetooth low energy (BLE), a variation of the Bluetooth standard, since these technologies are easy to set up with reasonable implementational cost [1], [8] as it makes use of available infrastructure unlike RFID and UWB which demands specialized hardware. Additionally, current mobile devices such as smart phones and laptops support Bluetooth and Wi-Fi [1], [3]. There are several works [9]–[11] that fuses these two technologies together to achieve complementary advantages [12].

Generally, the form of fingerprints that the system adopts are the received signal strength (RSS) which is the measurement of signal power received [13]. The indoor environment is complex as it contains various obstacles, e.g., wall and door, leading to non-line-of-sight propagation [14]. Wireless signal localization systems typically demand an accurate propagation model which can be difficult to established since the wireless signal fluctuates as it experiences diffraction, reflection and scattering [2] and does not follow the conventional path loss model. Even so, multiple studies have reported that the fingerprinting-based system produced good localization performance in complex indoor environment as compared to propagation modelling and geometric approach because the constructed radio map has taken into account the complicated signal patterns caused by reflection, shadowing and fading [15], [16]. Nevertheless, it cannot be claimed that the fingerprinting method is without drawbacks because the accuracy of the system is greatly influenced by the quality of radio map [17] and thus, each RP requires multiple signal samples collected from varying angle to ensure that the database is able to capture all the possible RSS variation. However, even with multiple RSS sample being collected, the dynamic indoor environments, where permanent and transient changes take place, could have a detrimental effect on the localization performance [18]. The structure, layout and presence of human can bring significant impact towards wireless signal [2]. Therefore, whenever an alteration happens to the surrounding, such as different placement of heavy objects, wireless signal will fluctuate [13]. According to earlier observations, fingerprinting expects the real-time data collected from the user device to closely matched with those in the training database obtained at the same position because a larger gap between those two data creates a higher positioning error [16]. This would mean that constant update of the database is required to maintain the accuracy of system which is an undesirable task because constructing a radio map is a time-consuming process and conducting a site survey can also be labour intensive [1], [2], [19].

Ultimately, it is critical to apply a robust learning algorithm that could assist in minimizing the positioning error of the system. In general, k-nearest neighbour (k-NN) algorithm has been the benchmark for more recent works. Extensive research has been conducted on the k-NN algorithm, as demonstrated in [20], which has shown satisfactory performance in positioning systems. However, numerous efforts have been made to further enhance the performance of these systems by minimizing the estimation error in positioning. Most of these studies have shifted their focus to deep learning rather than simple machine learning due to its great learning capability as deep learning algorithm is able to automatically perform feature extraction. A Wi-Fi fingerprint localization method using a four-layer deep neural network pre-trained by a stacked auto encoder for coarse localization was proposed [21]. Then, a hidden Markov model (HMM) localizer is used to further refine the initial position estimation. Kim *et al*. [22], the DNN-based positioning system follows a hierarchical approach for building and floor classification where Wi-Fi fingerprints were taken as input. Furthermore, the dimensionality of the input data was reduced by incorporating stacked auto encoder. The researchers [23], [24] constructed an image from received signal strength indicator (RSSI) fingerprint so that it could be fed into a 2-dimensional (2D) convolutional neural network (CNN) for indoor positioning whereas in [25], a 1-dimensional (1D) CNN was used. The DeepFi [26] and ConFi [27] positioning system also uses a CNN framework, however, the fingerprint information used are CSI which requires additional hardware.

The latest works in [28], [29] have integrated knowledge distillation to the CNN-based positioning system and the motivation behind it is to increase the positioning accuracy of a simple CNN network by leveraging knowledge distillation to acquire valuable knowledge possessed by a pre-trained complex CNN network. From the results shown in [28], [29], it has been proven that the knowledge distilled IPS is a great alternative to be used in a resource-constraint devices for real-time localization as compared to a complex system as it exhibits a high localization accuracy while maintaing lower processing and execution time. Although these knowledge distilled positioning techniques seem promising, it is still unknown if they would be greatly affected by variation of wireless signals caused by changes in the positioning environment.

Given the high probability of its real-world deployment in a dynamic setting, an ideal positioning system must be resilient to changes in its surroundings. To the best of our knowledge, there is no existing literature that has explored the effects of alterations in environmental layout on the localization performance using knowledge distilled positioning methods. In order to address the crucial research gap concerning the unavailability of the study related to the practicality of knowledge distilled positioning techniques in real-world dynamic environments, which is essential for ensuring seamless user experience in real-world scenarios, this paper aims to examine the robustness of various knowledge distilled based positioning techniques in the presence of a layout change in the target space, which may impact the RSSs of wireless signals. The presence of obstacles within the target space can result in attenuation and multipath fading issues, thereby inducing substantial signal fluctuation throughout both the offline and online phases. Specifically, the study focuses on two knowledge distilled positioning techniques, namely knowledge distilled based indoor positioning with and without a teacher assistant, and benchmarks their performance to those of the CNN models with different architectures and sizes. A comprehensive evaluation of various positioning systems considered is conducted using a real-world hybrid dataset with Wi-Fi and BLE fingerprints that encompasses a variety of multi-floor indoor layouts. The impacts of changes in layout on the performance of the systems are also thoroughly analysed. The findings of this study are highly significant as they provide valuable insights to designers and operators of IPS, empowering them to improve the system's resilience to layout changes and deliver consistent and precise location-based services.

The remainder of this paper is structured as follows. Section 2 expounds on the working principles of knowledge distilled CNN based positioning systems, both with and without teacher-assistant. In section 3, the experimental environment for the collection of the indoor positioning dataset is elucidated. Following this, section 4 delineates the configurations of the knowledge distilled CNN based positioning systems and the evaluation metrics employed for performance comparisons. In section 5, the results and discussions of findings are presented. Finally, section 6 summarizes the important findings, discusses their impacts, and suggests avenues for future research.

2. METHODS

In the first part of this section, the details on the CNN-based indoor localization scheme are presented. Then, the remaining part of this section offers a comprehensive explanation on the working principles of the knowledge distilled CNN-based IPS and the teacher-assistant knowledge distilled CNN-based IPS. The primary goal of integrating the knowledge distillation schemes to the CNN-based IPS is to create a precise positioning system with low complexity, allowing the system to be easily installed on devices with constrained resources.

2.1. Convolutional neural network-based positioning system

Based on the findings from authors [28], [29], the primary algorithm employed for location classification is a 2D CNN algorithm. The number of location class will be equal to the number of RPs established in the target space. Given an indoor area with *M* RPs, the total number of training samples will be noted as $N = \sum_{m=1}^{M} g_m$ where the total number of samples at the *m*th RP is represented by g_m. The system takes in the RSSI input denoted by $\{\boldsymbol{r}^n | n = 1,2, \dots, N\}$ in which $\boldsymbol{r}^n$ is the *n*th samples of RSSI vector as (1):

$$\boldsymbol{r}^n = [r_1^n, r_2^n, \dots, r_K^n] \tag{1}$$

where $r_k^n k = 1,2, \dots, K$ denotes the RSSI detected from the *k*th AP and *K* is the total number of available AP as well as its corresponding ground truth which is denoted by $\{\boldsymbol{y}^n | n = 1,2, \dots, N\}$. Since the system is using a 2D-CNN algorithm, the 1D RSSI vector needs to be reshaped into a square 2D fingerprint image $\boldsymbol{X}^n$ of size $Q_1 \times Q_1$. To ensure that the 1D RSSI vector can be converted to a square fingerprint image, $\boldsymbol{r}^n$ is padded with zeroes if $K \neq c^2$where c is an integer. After passing the 2D fingerprint image as the input, it will go through the several convolutional layers and pooling layers for feature extraction and then it will be flattened and sent to the dense network where each neuron in the layer is connected to all the neuron in the adjacent layer for classification. In the final layer of the fully connected network, a softmax activation function, formulated by (2), is applied to compute the probability of each class.

$$f_{softmax}(x_j) = \frac{e^{x_j}}{\sum_{l=1}^{L} e^{x_l}} \tag{2}$$

Where x_j and x_l, $l = 1,2, \dots, L$is the logit of the *j*th neuron and *l*th neuron, respectively, and L is the total number of neurons for the layer considered. The number of neurons in the last fully connected layer is set to the total number location classes and the output of this layer will be a vector of size $1 \times M$ containing the probability for each location class. The output vector is indicated as (3):

$$\hat{\boldsymbol{y}}^n = [\hat{y}_1^n, \hat{y}_2^n, \dots, \hat{y}_M^n] \quad (3)$$

Probability of each location class consist of numbers ranging from 0 to 1 and the total probability for all location classes is 1. The class exhibiting the highest probability among the location classes is assumed to be the final position of the target. The algorithm is optimized by reducing the cross-entropy loss between the predicted output and the ground truth through (4):

$$\begin{aligned} L_{\mathrm{CE}}\left(\boldsymbol{z}^n, \boldsymbol{y}^n\right) &= H\left(f_{\mathrm{softmax}}(\boldsymbol{z}^n), \boldsymbol{y}^n\right) \\ &= -\sum_{k=1}^{M} f_{\mathrm{softmax}}(z_k^n)\log(y_k^n) \end{aligned} \quad (4)$$

where is $H(\psi, \xi) = -\sum_{k=1}^{M} \psi_k log(\xi_k)$ written as the cross-entropy loss function and $\boldsymbol{z}^n = [z_1^n, z_2^n, \dots, z_M^n]$ is the vector of logits produced at the final fully connected layer for the *n*th input sample.

2.2. Knowledge distilled convolutional neural network-based positioning system

According to the assertions made in [28], [29], it has been noted that although CNNs demonstrate impressive learning capabilities, the networks necessary to achieve satisfactory localization often exhibit high complexity, making them unsuitable for deployment on edge computing systems. To facilitate the operation of the system on resource-constrained devices, a knowledge distillation process is employed. This enables the positioning system to utilize an algorithm with reduced complexity, while still preserving a high level of positioning accuracy. Knowledge distillation is a concept of model compression that was popularized by [30] and the works in [28], [29] have applied the fundamental principle of knowledge distillation to enable a simple CNN algorithm to provide greater localization accuracy by relying on the informative dark knowledge from another CNN network with higher complexity. Evidently, two CNN models of varying complexity, i.e., the basic model being referred to as the student model and the pre-trained complex model known as the teacher model, are needed to accomplish knowledge distillation. Insightful knowledge that the student model obtained from the teacher model includes the softened probabilities of the teacher network rather than the hard prediction provided by the softmax activation function. Using (2), the output probabilities for all classes, except the correct class, will be relatively low, while the correct class will exhibit a significantly high probability. From there, not much information can be obtained to train the student network. Hence, logits of the teacher model must be scaled by a temperature parameter to softened the probability distribution, resulting a distinguishable inter-class relationship, using higher temperature. Firstly, the knowledge distilled positioning system operates by inputting the 2D fingerprint images to the pre-optimized teacher network which will then map out those images to logit vector $\boldsymbol{z} = [z_1, z_2, \dots, z_M]$. After that, soft labels are created by applying a temperature-scaled softmax activation function to the aforementioned logits. The temperature-scaled softmax activation function by (5):

$$\rho_i = f_{TS-Softmax}(z_i) = \frac{e^{\frac{z_i}{T}}}{\sum_{j=1}^{M} e^{\frac{z_j}{T}}} \quad (5)$$

where the temperature parameter is indicated by $T \geq 1$. It is recorded that when T=1, $f_{TS-Softmax}(z_i) = f_{softmax}(z_j)$.

At the same time, the 2D fingerprint image is also fed to the student network which will generate the hard and soft outputs using (2) and (5), respectively. The student model is trained using an altered loss function formulated as (6):

$$L = \alpha L_{CE}(\boldsymbol{z}_s^n, \boldsymbol{y}^n) + \beta L_{KD}(\boldsymbol{z}_s^n, \boldsymbol{z}_t^n) \quad (6)$$

where $L_{CE}(\boldsymbol{z}_s^n, \boldsymbol{y}^n)$ represents the cross-entropy loss between predicted output of the student network and its ground truth, and $L_{KD}(\boldsymbol{z}_s^n, \boldsymbol{z}_t^n)$ is the distillation loss. The significance of $L_{CE}(\boldsymbol{z}_s^n, \boldsymbol{y}^n)$ and $L_{KD}(\boldsymbol{z}_s^n, \boldsymbol{z}_t^n)$ is represented by α and β, respectively, where is $\alpha = 1 - \beta$ and $\alpha \in [0,1]$. A higher value of α indicates that the student is trained in a manner that is following more closely to the student loss than it is to the distillation loss and vice versa. The distillation loss is written as (7):

$$L_{KD}(\boldsymbol{z}_s^n, \boldsymbol{z}_t^n) = T^2 D_{KL}(f_{TS-Softmax}(\boldsymbol{z}_s^n), f_{TS-Softmax}(\boldsymbol{z}_t^n)) \quad (7)$$

where $D_{KL}(\psi_k, \xi_k) = \sum_{k=1}^{M} \psi_k log\left(\frac{\psi_k}{\xi_k}\right)$ is the Kullback-Leibler (KL) divergence, $\boldsymbol{z}_s^n$ and $\boldsymbol{z}_t^n$ are the logit vector of the *n*th sample generated by the student and the teacher network, respectively. The KL divergence function is applied to assess the divergence between the softened probability distribution of the student network and the teacher network.

2.3. Teacher-assistant knowledge distilled convolutional neural network-based positioning system

It has been demonstrated that a positioning system trained on a basic student model can accomplish a low localization error comparable to the localization error of a positioning system trained on a large teacher model through the application of knowledge distillation. Nevertheless, using knowledge distillation might not always be beneficial, particularly if there is a substantial complexity gap between the student network and the teacher network employed to supervise the student network. Given the large disparity in complexity between the simple student model and the cumbersome teacher model, the student model may not posses the capacity to replicate the performance of the teacher network. Furthermore, a highly complex teacher model is more confident in its predictions and thus, making the logits less soft, reducing the effectiveness of knowledge distillation. Therefore, a solution to bridge the teacher-student network gap is to introduce a teacher-assistant network in the indoor positioning scheme [29]. It is crucial to note that the teacher-assistant network will act as an intermediary network and hence, the size and complexity of the network must lie between that of the student network and the teacher network.

The working principle of the teacher-assistant knowledge distilled CNN-based positioning system is quite similar to the working principle of knowledge distilled CNN-based positioning system discussed in section 2.2. However, the student network does not directly receive the knowledge from the teacher network. Instead, the information passes through the teacher-assistant network first and only then, the teacher-assistant network will pass down the useful knowledge to the student network. During the training of the teacher-assistant network, the cost function applied by the teacher-assistant network can be formulated as (8):

$$L^{TA-TAKD} = \alpha L_{CE}(\boldsymbol{z}_{ta}^n, \boldsymbol{y}^n) + \beta L_{KD}(\boldsymbol{z}_{ta}^n, \boldsymbol{z}_t^n) \quad (8)$$

where $L_{CE}(\boldsymbol{z}_{ta}^n, \boldsymbol{y}^n)$ represents the cross-entropy loss between predicted output of the teacher-assistant network and its ground truth, and $L_{KD}(\boldsymbol{z}_{ta}^n, \boldsymbol{z}_t^n)$ is the distillation loss. The significance of $L_{CE}(\boldsymbol{z}_{ta}^n, \boldsymbol{y}^n)$ and $L_{KD}(\boldsymbol{z}_{ta}^n, \boldsymbol{z}_t^n)$ is represented by α and β, respectively, where is $\alpha = 1 - \beta$ and $\alpha \in [0,1]$. The distillation loss between the teacher-assistant network and teacher network is expressed as (9):

$$L_{KD}(\boldsymbol{z}_{ta}^n, \boldsymbol{z}_t^n) = T^2 D_{KL}(f_{TS-Softmax}(\boldsymbol{z}_{ta}^n), f_{TS-Softmax}(\boldsymbol{z}_t^n)) \quad (9)$$

where $\boldsymbol{z}_{ta}^{\,n}$ and $\boldsymbol{z}_t^n$ are the logit vector of the *n*th sample generated by the teacher-assistant network and the teacher network, respectively.

Subsequently, the student network is trained by leveraging the information provided by the ground truth and the information acquired from the soft logits of both the student and teacher networks. The following loss function is utilized when training the student network as (10):

$$L^{S-TAKD} = \alpha L_{CE}(\boldsymbol{z}_s^n, \boldsymbol{y}^n) + \beta L_{KD}(\boldsymbol{z}_s^n, \boldsymbol{z}_{ta}^n) \quad (10)$$

where $L_{CE}(\boldsymbol{z}_s^n, \boldsymbol{y}^n)$ represents the cross-entropy loss between predicted output of the student network and its ground truth, and $L_{KD}(\boldsymbol{z}_s^n, \boldsymbol{z}_{ta}^n)$ is the distillation loss. The significance of $L_{CE}(\boldsymbol{z}_s^n, \boldsymbol{y}^n)$ and $L_{KD}(\boldsymbol{z}_s^n, \boldsymbol{z}_{ta}^n)$ is represented by α and β, respectively, where is $\alpha = 1 - \beta$ and $\alpha \in [0,1]$. The distillation loss between the student network and teacher-assistant network is written as (11):

$$L_{KD}(\boldsymbol{z}_s^n, \boldsymbol{z}_{ta}^n) = T^2 D_{KL}(f_{TS-Softmax}(\boldsymbol{z}_s^n), f_{TS-Softmax}(\boldsymbol{z}_{ta}^n)) \quad (11)$$

where $\boldsymbol{z}_s^n$ and $\boldsymbol{z}_{ta}^n$ are the logit vector of the *n*th sample generated by the student and the teacher-assistant network, respectively.

3. EXPERIMENTAL ENVIRONMENT

In real-world scenario, the indoor environment is dynamic and wireless signal tends to fluctuate with changes that occur in the indoor space, i.e., change in furniture placement, addition of obstacles and object movements. Since the aim of this work is to study the feasibility of the knowledge distilled positioning schemes in dynamic indoor environments, the dataset required must consist of the RSS map gathered in the same space with at least two different set ups. Therefore, we have decided to utilize the hybrid-fingerprint data layout change (HDLC) dataset [31] because this dataset provides training and testing data with different layouts. The layout used to train the positioning systems is without any obstacles and it will be referred as the original layout. In order to evaluate the performance of the positioning systems, the considered systems will be tested with the original layout as well as another layout that contains obstructions, referred to as altered layout. The obstructions comprise partition boards with 1.8 m in length, 1.5 m in width, and 0.0127 m in thickness.

The measurement campaign for this dataset was conducted in the Faculty of Engineering (FOE), Multimedia University, Cyberjaya and calibration point was set on the ground, first and second floor of Wing C of the FOE building. There are 96 RPs established in the ground floor and 144 RP each in the first and second floor, resulting in a total of 384 RPs altogether. Note that all of the RPs are evenly spaced, at a distance of 1 m. To ensure that a full range of delectable RSSI were captured, 30 samples were gathered at each RP, whereby 20 samples were stored as training data and the remaining 10 samples being stored as testing data. There are 42 Wi-Fi APs and 17 BLE beacons used to transmit signals. The RSSI readings from the Wi-Fi APs and BLE beacons are stored along with the floor level, *x*-coordinate and *y*-coordinate of RPs, resulting in a database with 62 attributes. The entire floor plan of the indoor space along with the arrangement of partition boards is illustrated in Figure 1. Figures 1(a) to (c) are the floor plans for the ground floor, first floor and second floor, respectively.

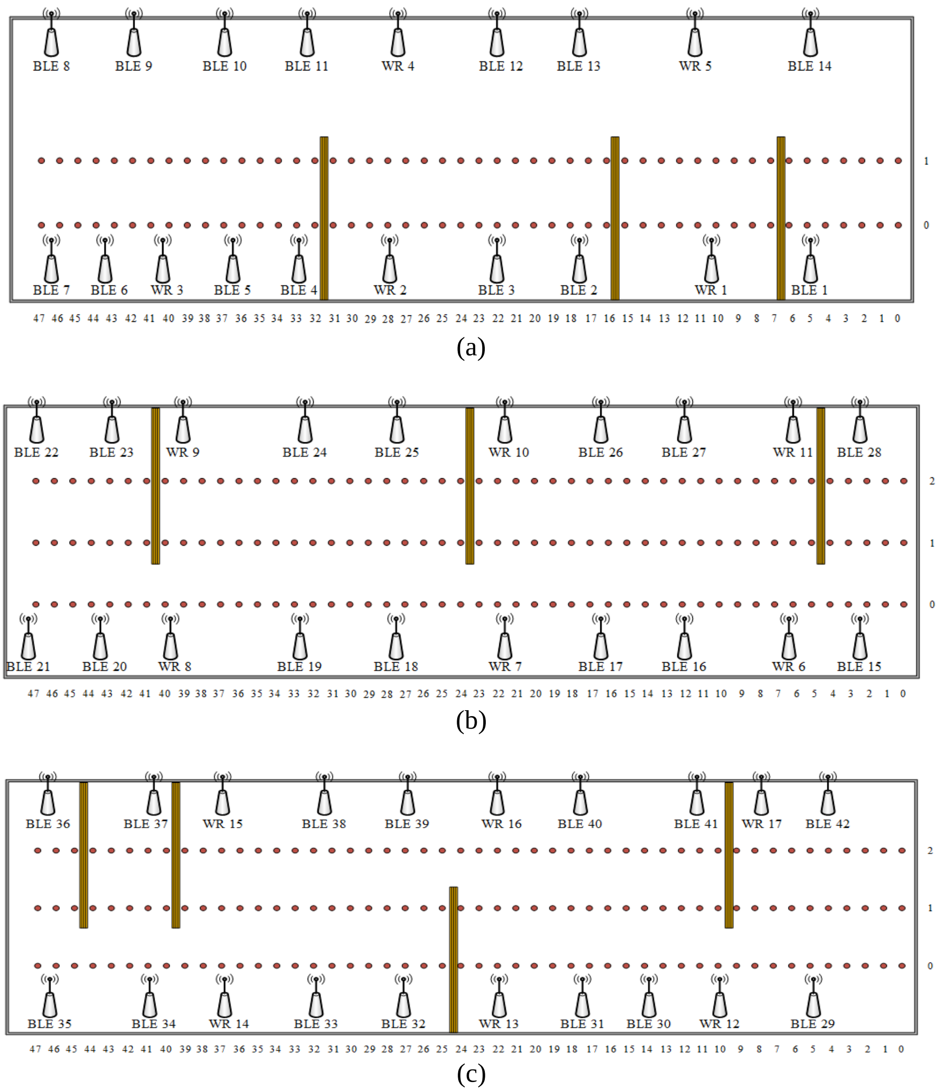

Figure 1. Floor plan of the experimental environments with partition board; (a) ground floor, (b) first floor, and (c) second floor

4. SIMULATION SETUP
4.1. Convolutional neural network model configuration

For this study, three CNN models were established, labelled as the teacher model, student model and the teacher assistant model. Python 3.7.12 is utilized to carry out the simulations and Keras 2.7.0 is used to established the deep learning models. The configuration of the three models is tabulated in the Table 1. The models are presented in the order of decreasing complexity: teacher model, teacher-assistant model, and student model; whereby the teacher model, teacher-assistant model and student model consist of six convolutional layers, four convolutional layers and one convolutional layer, respectively.

Table 1. Configuration of models considered

CNN model	Settings
Teacher	No of convolutional layers: 6 Filter size: 2×2 No of filters: 4,4,8,8,16,16 Activation function after convolutional layers: ReLU No of max pooling layers: 3 Kernel size: 2×2 Strides: 1×1 Hidden layer: 500 (dropout=0.3) Output node: 384 nodes
Teacher-assistant	No of convolutional layers: 4 Filter size: 2×2 No of filters: 4,4,16,16 Activation function after convolutional layers: ReLU No of max pooling layers: 2 Kernel size: 2×2 Strides: 1×1 Hidden layer: 500 (dropout=0.3) Output node: 384 nodes
Student	No of convolutional layers: 1 Filter size: 2×2 No of filters: 32 Activation function after convolutional layers: ReLU No of max pooling layers: 1 Kernel size: 2×2 Strides: 1×1 Output node: 384 nodes

Four positioning schemes, namely CNN-IPS (TM), CNN-IPS (SM), KD-CNN-IPS and TAKD-CNN-IPS, were developed by utilizing either a single basic model or a combination of the basic models. More specifically, the first scheme is the CNN-IPS (TM) and it solely employs the teacher CNN model. The second scheme is the CNN-IPS (SM) and it comprises only the student model. These two schemes, CNN-IPS (TM) and CNN-IPS (SM) serve as the baseline models and are expected to exhibit a trade-off between the positioning accuracy and execution time. On the other hand, the other two schemes incorporate knowledge distillation. KD-CNN-IPS framework is designed by having the student model trained under the supervision of the teacher model so that the student model is able to exploit the dark knowledge from the teacher model. Lastly, TAKD-CNN-IPS utilizes all the three models and the knowledge is transferred from the teacher model to the student model by passing it through the intermediary teacher-assistant model. Table 2 provides a clearer depiction of models utilized and the training process for each of the positioning schemes considered in this study. To simulate layout change, the positioning systems are trained and tested with datasets that are collected under two distinct arrangements for the same indoor space. The two different layouts can be produced by introducing obstacles, such as divider boards. As mentioned in section 3, the positioning schemes are initially trained using the original layout dataset. Subsequently, to assess their practicality in dynamic environments, they are tested using the altered layout dataset.

Table 2. Models utilized in the positioning schemes

Positioning scheme	Models
CNN-IPS (TM)	Teacher
CNN-IPS (SM)	Student
KD-CNN-IPS	Teacher → student
TAKD-CNN-IPS	Teacher → teacher-assistant → student

4.2. Performance metrics

The performance metrics used to measure the practicality of the positioning schemes in this work include location class accuracy, floor accuracy, average positioning error and execution time. The location class accuracy γ_C and floor accuracy $\gamma_F$are defined by (12) and (13), respectively:

$$\gamma_C = \frac{N_C}{N_T} \times 100\% \tag{12}$$

$$\gamma_F = \frac{N_F}{N_T} \times 100\% \tag{13}$$

where N_C is the number of correctly predicted class samples, $N_F$is the number of correctly predicted floor samples and N_T is the total number of test samples. As for the average errors, both 3D and 2D positioning errors are considered to provide a comprehensive analysis and they are expressed by (14) and (15), respectively:

$$e_{3D} = \frac{1}{N_T} \sum_{n=1}^{N_T} \sqrt{(\tilde{x}_n - \hat{x}_n)^2 + (\tilde{y}_n - \hat{y}_n)^2 + (\tilde{z}_n - \hat{z}_n)^2} \tag{14}$$

$$e_{2D} = \frac{1}{N_F} \sum_{f=1}^{N_F} \sqrt{(\tilde{x}_f - \hat{x}_f)^2 + (\tilde{y}_f - \hat{y}_f)^2} \tag{15}$$

where $(\tilde{x}_n, \tilde{y}_n, \tilde{z}_n)$ and $(\hat{x}_n, \hat{y}_n, \hat{z}_n)$ are the *i*th true and predicted coordinate, respectively.

To provide a deeper insight into the impact of indoor layout change on the systems' localization performance, the degradation in location class accuracy due to the layout change is calculated as (16):

$$\Delta\gamma_C = \gamma_C^O - \gamma_C^A \tag{16}$$

where γ_C^O and γ_C^A represent the location accuracy of the original layout and altered layout, respectively. Intuitively, a positive value of $\Delta\gamma_C$ implies a reduction in accuracy and vice versa.

Since the primary goal of this study is to examine whether the knowledge distilled frameworks can still improve the localization performance of a simple CNN positioning scheme after the wireless conditions of the indoor space has altered, two evaluation metrics, which are the performance gains of the systems against the CNN-IPS (SM) in terms of 3D average positioning error P_{3D} and the performance gains of the systems against the 2D average positioning errors, are introduced. The performance gains are formulated as (17):

$$P_\varepsilon = \frac{e_\varepsilon^B - e_\varepsilon^W}{e_\varepsilon^B} \times 100\% \tag{17}$$

where e_ε^B is the average positioning error of baseline CNN-IPS (SM) and e_ε^W is the average positioning error of the system considered; P_ε represents the 3D performance gain if $\varepsilon = 3D$, whereas it signifies the 2D performance gain if $\varepsilon = 2D$.

5. RESULTS AND DISCUSSION

This section investigates the impact of layout change on the positioning performance of KD-CNN-IPS and TAKD-CNN-IPS. It is worth highlighting that these knowledge distilled schemes are supposed to produce better positioning accuracy than a CNN network of the same complexity and possess a shorter executing time than a CNN model with higher complexity. Thus, to investigate whether the schemes are working as intended in the given circumstances, a thorough analysis is conducted by benchmarking the performance of the two schemes against both CNN-IPS (TM) and CNN-IPS (SM). Table 3 tabulates the location class accuracy, training time and testing time of the CNN-IPS (TM) and CNN-IPS (SM). During the training phase, both schemes are trained using the data from the original layout. As anticipated, the results indicate that CNN-IPS (TM) achieves a higher training accuracy compared to CNN-IPS (SM) as its network incorporates more convolutional layers. However, this improvement in accuracy comes at the expense of longer training time. Specifically, CNN-IPS (TM) exhibits a 28.61% higher training accuracy than CNN-IPS (SM), but the training time is increased by 87.34% as a trade-off. In the testing phase, where the two schemes are evaluated using the testing data of the original layout and the altered layout, it is observed that CNN-IPS (TM) outperforms CNN-IPS (SM) in terms of testing accuracy for both layouts. Additionally, the testing time of CNN-IPS (TM) is also higher than that of the CNN-IPS (SM) for both layouts. These observations are consistent with the results obtained during the training phase.

Table 3. The classification performance of CNN-IPS (TM) and CNN-IPS (SM)

Technique	Training phase		Testing phase			
			Original layout		Altered layout	
	γ_C (%)	Time (s)	γ_C (%)	Time (s)	γ_C (%)	Time (s)
CNN-IPS (TM)	35.52	51.47579	12.55	1.4271	4.77	1.3119
CNN-IPS (SM)	6.91	6.5139	6.35	0.6959	3.31	0.7181

The results in Table 3 clearly demonstrate that the size and structure of CNN play a crucial role in ensuring good localization performance. Generally, a larger and more complex network leads to a more accurate positioning system at the price of increased training and execution time. This is because a more complex model typically possesses an enhanced capability to decipher the underlying complex nonlinear relationship between the RSSIs and the user locations. However, the complexity of such model may pose a significant challenge when it comes to practical implementation on resource-constrained devices. To address the inherent trade-off between localization performance and execution time, the knowledge distilled positioning frameworks were designed. These frameworks aim to strike a balance by improving the classification accuracy and reducing the positioning error of a simple CNN-based localization system while maintaining the execution time of the system. Another important observation from Table 3 is that the positioning accuracies for both the CNN-IPS (TM) and CNN-IPS (SM) degrade when tested with the altered layout. When the CNN-IPS (TM) and CNN-IPS (SM) are assessed using the data from the altered layout, the positioning accuracy of CNN-IPS (TM) and CNN-IPS (SM) decreases by 7.78% and 3.04%, respectively. This performance degradation is due to the presence of partition boards that obstruct certain APs, resulting in signal transmission being influenced by reflections and fading. Consequently, the distribution of signal vector received at each point of interest might deviate from those stored in the database. As mentioned previously, the accuracy of a fingerprint-based localization system heavily relies on the database. Hence, any changes to the layout of the indoor space will inevitably result in a deterioration of the positioning systems' performance.

In this work, more emphasis is placed on the localization performance of the knowledge distilled positioning schemes as the primary objective is to examine the robustness of these schemes to changes in the indoor layout. Figure 2 shows the degradation in location class accuracy caused by the layout change for CNN-IPS (TM), CNN-IPS (SM), KD-CNN-IPS, and TAKD-CNN-IPS. As expected, all schemes are negatively affected by the layout change. Based on Figure 2, CNN-IPS (TM) suffers the highest degradation in location class accuracy from the layout change. It is also observed that the impact of layout changes on CNN-IPS (SM) and KD-CNN-IPS are almost similar. More specifically, when the indoor layout varies, CNN-IPS (SM) suffers a performance degradation in positioning accuracy of 3.04%, while the positioning accuracy of KD-CNN-IPS drops by 3.05%. In comparison to CNN-IPS (SM) and KD-CNN-IPS, the performance loss due to layout change is more severe for TAKD-CNN-IPS, with a performance loss of 3.44%. The reason for these performance trends can be elucidated as follows. When the layout of the indoor environment changes, the trained teacher model's ability to capture the relationship between RSSIs and locations for the new environment deteriorates. This issue is exacerbated when TAKD-CNN-IPS is adopted as it involves an additional network than KD-CNN-IPS, i.e., teaching assistant. The modeling errors will be propagated from the teacher model to the teacher assistant and subsequently to student model, which will further degrade the RSSs and location relationship modeling capability. As a result, TAKD-CNN-IPS experiences a more pronounced decrease in performance due to the layout change.

Aside from the location class accuracy, it is important to examine the performance of other evaluation metrics such as the positioning error and the floor accuracy as these metrics can provide valuable insight to the study. Table 4 presents the positioning performance of different schemes considered during the testing phase, encompassing floor accuracy, 2D and 3D average positioning errors for both the original and altered layouts. From the Table 4, it is evident that when the layouts of the training and testing phases are the same, all schemes achieve 100% accuracy in predicting the correct floor. Since the floor accuracies for all schemes are 100%, the 2D and 3D average positioning errors are identical. Both the 2D and 3D average positioning errors follow a decreasing trend in the sequence of CNN-IPS (SM), KD-CNN-IPS, TAKD-CNN-IPS, and CNN-IPS (TM). The performance gains of CNN-IPS (TM), KD-CNN-IPS and TAKD-CNN-IPS over the CNN-IPS (SM) in terms of 3D and 2D average positioning error during the testing phase are illustrated in Figures 3 and 4, respectively. Notably, the incorporation of knowledge distillation in KD-CNN-IPS and TAKD-CNN-IPS lead to an improvement in the positioning error of the baseline CNN-IPS (SM) by 3.46% and 10.04%, respectively. The superior performance of KD-CNN-IPS and TAKD-CNN-IPS over CNN-IPS (SM) can be attributed to their ability to retain the high-quality performance of the complex teacher model through the utilization of softened probabilities acquired from the teacher network. By harnessing these softened probabilities, which encapsulates richer inter-class relations, the learning process is significantly enhanced. Furthermore, the reason that TAKD-CNN-IPS outperforms its KD-CNN-IPS counterpart in this scenario is due to the utilization of an intermediate network as teacher assistant. This enables the student model of TAKD-CNN-IPS to acquire a greater amount of knowledge by being trained with the soft distribution generated by a network that possesses a closer capacity, resulting in lower distillation and student losses.

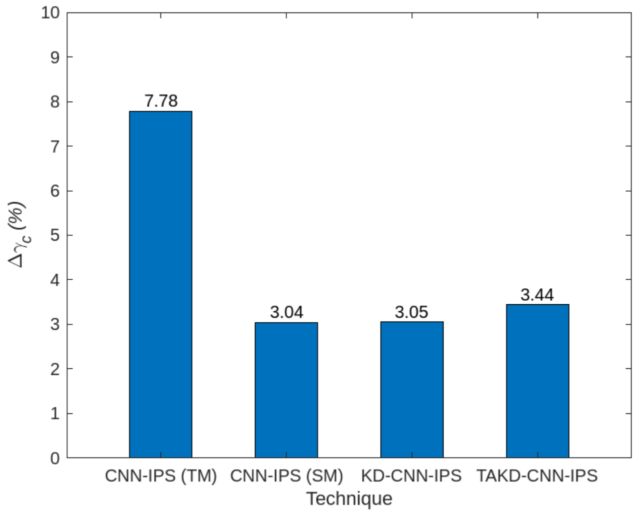

Figure 2. The loss in testing accuracy performance for different technique considered resulting from modifications in the indoor layout

Table 4. The floor accuracy and average positioning error evaluated for both the original and altered layouts during the testing phase

Technique	Original layout			Altered layout		
	3D average positioning error (m)	Floor accuracy (%)	2D average positioning error (m)	3D average positioning error (m)	Floor accuracy (%)	2D average positioning error (m)
CNN-IPS (TM)	2.5647	100	2.5647	7.9133	78.67	5.6052
CNN-IPS (SM)	3.4078	100	3.4078	10.1980	79.48	6.1182
KD-CNN-IPS	3.2900	100	3.290	9.0192	82.89	8.0509
TAKD-CNN IPS	3.0655	100	3.0655	11.6409	75.47	8.0739

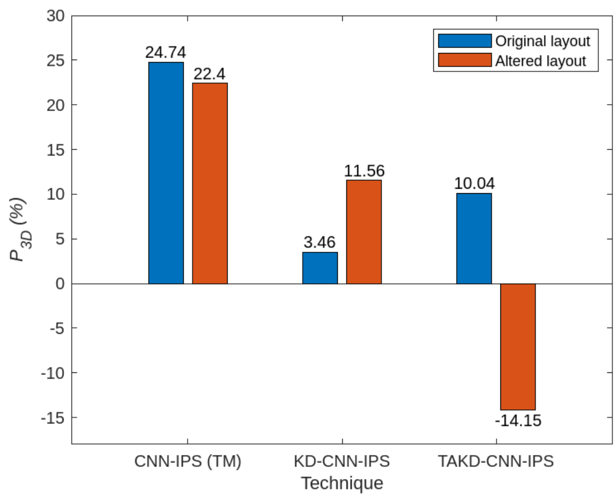

Figure 3. The performance gains of CNN-IPS (TM), KD-CNN-IPS, and TAKD-CNN-IPS over the CNN-IPS (SM) in terms of 3D average positioning error during the testing phase

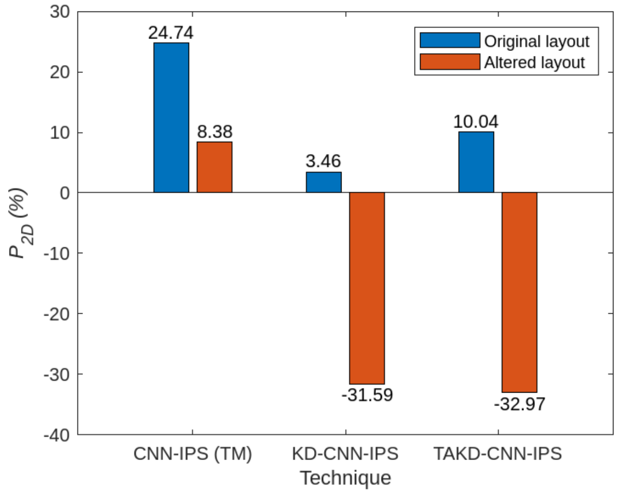

Figure 4. The performance gains of CNN-IPS (TM), KD-CNN-IPS, and TAKD-CNN-IPS over the CNN-IPS (SM) in terms of 2D average positioning error during the testing phase

When the indoor layout is altered, it can be observed that the floor accuracies and the average positioning errors for all techniques degrade. More explicitly, KD-CNN-IPS demonstrates the highest floor accuracy of 82.89%, followed by CNN-IPS (SM), CNN-IPS (TM), and TAKD-CNN-IPS with floor accuracies of 79.48%, 78.67%, and 75.47%, respectively. From the findings, it is evident that TAKD-CNN-IPS suffers the highest degradation in terms of floor accuracy, followed by CNN-IPS (TM) and CNN-IPS (SM), while KD-CNN-IPS experiences the smallest degradation. These findings indicate that although the CNN-IPS (TM) has the best performance without any layout change, its performance is more susceptible to the layout change as compared to the KD-CNN-IPS. Although CNN-IPS (TM) does not have the best floor accuracy among the positioning schemes under consideration, it still provides the lowest 2D and 3D positioning errors. Nevertheless, CNN-IPS (TM) shows a large performance degradation in terms of average positioning error, specifically the 3D average positioning error. More specifically, CNN-IPS (TM) still demonstrates a higher degradation of 208.55% as compared to KD-CNN-IPS and CNN-IPS (SM), which degrades by 174.13% and 199.25%, respectively. This further substantiates that, in comparison to the KD-CNN-IPS and CNN-IPS (SM), the CNN-IPS (TM) positioning scheme is more sensitive to the environmental dynamics. It also implies that under more severe changes in the indoor environment, CNN-IPS (TM) may perform worse than the KD-CNN-IPS scheme.

From the performance gains of CNN-IPS (TM), KD-CNN-IPS and TAKD-CNN-IPS over the CNN-IPS (SM) in terms of 3D and 2D average positioning error depicted in Figures 3 and 4, another observation that can be seen is that the KD-CNN-IPS framework exhibits a lower 3D average positioning error than CNN-IPS (SM), but it has a higher 2D average positioning error. Considering that the wrongly predicted floor is discarded in the calculation of 2D average positioning error, it is safe to presume that most samples which CNN-IPS (SM) provided an incorrect floor prediction carry a significant distance error. Ultimately, this implies that knowledge distillation implemented in the KD-CNN-IPS still has the ability to improve the positioning performance of the baseline CNN-IPS (SM). Conversely, the knowledge distillation framework implemented in the TAKD-CNN-IPS fails to bring any improvement over the CNN-IPS (SM). In fact, both the floor accuracy and average positioning performance of TAKD-CNN-IPS are worse than those of the CNN-IPS (SM). When the indoor layout changes, wireless signals at every RP are affected differently, depending on the adjustments made. As a result, the wireless signal distribution could be entirely different from the radio map. It is well known that the performance of fingerprint-based localization depends greatly on the radio map. Higher error counts may result from a wide disparity between the target RSSI vector and the radio map. Evidently, the training process of the TAKD-CNN-IPS occurs sequentially. The student model in the positioning scheme learns from teacher-assistant model that is pre-trained by a teacher model. If an error is made during the learning process of the teacher model, the same error will undoubtedly be passed down to both the teacher-assistant model and student model as proclaimed in [32], leading to a substantial error accumulation. This argument would justify the deterioration in positioning accuracy of TAKD-CNN-IPS.

As mentioned previously, the knowledge distilled positioning schemes are designed to reduce the complexity of deep learning techniques. In comparison to the conventional deep learning strategies, knowledge distilled positioning approaches are more practical to be deployed on devices with limited resources. This is due to their reduced computational requirements, resulting faster positioning times. Such efficiency is particularly crucial for real-time positioning systems to ensure an excellent user experience. Table 5 provides insights into the execution time of the different positioning techniques examined. The results in Table 5 reveal that KD-CNN-IPS and TAKD-CNN-IPS have almost similar execution times to the CNN-IPS (SM) for both the original and altered layouts. This is because only the student model of KD-CNN-IPS and TAKD-CNN-IPS are executed during the online localization phase. Since the student model of KD-CNN-IPS and TAKD-CNN-IPS have the same architecture as CNN-IPS (SM), the execution times of all three positioning systems will be similar. Furthermore, due to the substantially simpler architecture of the student model employed in CNN-IPS (SM), KD-CNN-IPS, and TAKD-CNN-IPS compared to CNN-IPS (TM), the execution times for all three systems for both the original and modified layouts are only approximately 50% of those required by CNN-IPS (TM). This solidifies the effectiveness of knowledge distilled frameworks in minimizing the execution time, which is vital for real-time positioning systems to deliver a seamless user experience.

Table 5. Execution time of different positioning schemes tested on original and altered layouts during the testing phase

Technique	Execution time (s)	
	Original layout	Altered layout
CNN-IPS (TM)	1.4271	1.3119
CNN-IPS (SM)	0.7181	0.6959
KD-CNN-IPS	0.7000	0.6866
TAKD-CNN-IPS	0.7157	0.6857

6. CONCLUSION

This paper investigates the robustness of knowledge distilled frameworks, specifically KD-CNN-IPS and TAKD-CNN-IPS, to layout changes and obstacles that could obstruct the wireless signal propagation, resulting in attenuation and multipath fading issues. To assess their robustness, datasets with different training and testing environments are utilized. The knowledge distilled frameworks rely on a pre-trained network known as the teacher network to transfer the informative knowledge. The teacher network indicated by CNN-IPS (TM) along with CNN-IPS (SM), which is a positioning system with the same architecture as the student model of KD-CNN-IPS and TAKD-CNN-IPS, are used as baseline models for the knowledge distilled framework.

Our results reveal that in a static indoor environment, both the knowledge distilled frameworks exhibit superior positioning accuracy than CNN-IPS (SM) and maintain similar execution time as CNN-IPS (SM). In addition, their model complexities are lower and their execution times are shorter compared to those of CNN-IPS (TM), rendering them appealing solutions for deployment on resource-constrained devices. However, when the layout of the indoor environment changes, although all the knowledge distilled schemes are still able to maintain the execution time of CNN-IPS (SM), only KD-CNN-IPS manage to achieve a better localization performance than CNN-IPS (SM). Conversely, due to the more severe propagation of RSSIs and location modeling errors incurred by the teacher assistant in TAKD-CNN-IPS, TAKD-CNN-IPS fails to perform as anticipated when the layout of the indoor environment is altered, resulting in a 14.15% degradation in average positioning error. Aside from that, our results demonstrate that the KD-CNN-IPS is less vulnerable to layout change compared to other schemes, as evidenced by its minimal degradation when the indoor layout changes. Consequently, KD-CNN-IPS is a preferable choice for dynamic environments compared to TAKD-CNN-IPS, whereas TAKD-CNN-IPS is more suitable for static environments.

The insights derived from this work are highly valuable for both the research community and industry. They empower designers and operators of IPS to enhance the system's robustness to layout changes and consistently deliver precise real-time location-based services on resource constrained devices, such as edge devices, smartphones and embedded sensor nodes. These location-based services encompass a wide range of applications, which include but are not limited to positioning, tracking, and navigation. Furthermore, the insights uncovered in this study hold significant implications across diverse domains, such as activity recognition, facial recognition, and autonomous drone and self-driving car technologies. Future research endeavors may explore the integration of lifelong learning mechanisms into KD-CNN-IPS and TAKD-CNN-IPS to enhance their robustness in dynamic environments.

ACKNOWLEDGEMENTS

The authors acknowledge the support provided by the Ministry of Higher Education Malaysia in funding the study. This work was supported by the Fundamental Research Grant Scheme (FRGS), Ministry of Higher Education Malaysia under grant FRGS/1/2023/ICT02/MMU/02/7.

REFERENCES

[1] A. K. M. M. Hossain and W. S. Soh, "A survey of calibration-free indoor positioning systems," *Computer Communications*, vol. 66, pp. 1–13, Jul. 2015, doi: 10.1016/j.comcom.2015.03.001.
[2] I. H. Alshami, N. A. Ahmad, S. Sahibuddin, and F. Firdaus, "Adaptive indoor positioning model based on WLAN-fingerprinting for dynamic and multi-floor environments," *Sensors (Switzerland)*, vol. 17, no. 8, p. 1789, Aug. 2017, doi: 10.3390/s17081789.
[3] C. H. Hsieh, J. Y. Chen, and B. H. Nien, "Deep Learning-Based Indoor Localization Using Received Signal Strength and Channel State Information," *IEEE Access*, vol. 7, pp. 33256–33267, 2019, doi: 10.1109/ACCESS.2019.2903487.
[4] M. Uradzinski, H. Guo, X. Liu, and M. Yu, "Advanced Indoor Positioning Using Zigbee Wireless Technology," *Wireless Personal Communications*, vol. 97, no. 4, pp. 6509–6518, Dec. 2017, doi: 10.1007/s11277-017-4852-5.
[5] Z. Dong, C. Mengjiao, and L. Wenjuan, "Implementation of indoor fingerprint positioning based on ZigBee," in *Proceedings of the 29th Chinese Control and Decision Conference, CCDC 2017*, May 2017, pp. 2654–2659, doi: 10.1109/CCDC.2017.7978963.
[6] M. Anjum, M. A. Khan, S. A. Hassan, A. Mahmood, and M. Gidlund, "Analysis of RSSI fingerprinting in LoRa networks," in *2019 15th International Wireless Communications and Mobile Computing Conference, IWCMC 2019*, Jun. 2019, pp. 1178–1183, doi: 10.1109/IWCMC.2019.8766468.
[7] A. Poulose and D. S. Han, "UWB indoor localization using deep learning LSTM networks," *Applied Sciences (Switzerland)*, vol. 10, no. 18, p. 6290, Sep. 2020, doi: 10.3390/APP10186290.
[8] I. N. Kosović and T. Jagušt, "Enhanced Weighted Centroid Localization Algorithm for Indoor Environments," *International Conference on Communication Networks and Applications (ICCNA 2014)*, vol. 8, no. 7, pp. 1184–1188, 2014.
[9] L. Kanaris, A. Kokkinis, A. Liotta, and S. Stavrou, "Fusing Bluetooth beacon data with Wi-Fi radiomaps for improved indoor localization," *Sensors (Switzerland)*, vol. 17, no. 4, p. 812, Apr. 2017, doi: 10.3390/s17040812.
[10] Y. Miyashita, M. Oura, J. F. De Paz, K. Matsui, G. Villarrubia, and J. M. Corchado, "Preliminary study of classifier fusion based indoor positioning method," in *Advances in Intelligent Systems and Computing*, vol. 476, 2016, pp. 161–166, doi: 10.1007/978-3-319-40114-0_18.
[11] J. H. Seong and D. H. Seo, "Environment Adaptive Localization Method Using Wi-Fi and Bluetooth Low Energy," *Wireless Personal Communications*, vol. 99, no. 2, pp. 765–778, Mar. 2018, doi: 10.1007/s11277-017-5151-x.
[12] F. Liu *et al.*, "Survey on WiFi-based indoor positioning techniques," *IET Communications*, vol. 14, no. 9, pp. 1372–1383, Jun. 2020, doi: 10.1049/iet-com.2019.1059.
[13] P. Roy and C. Chowdhury, "A survey on ubiquitous WiFi-based indoor localization system for smartphone users from implementation perspectives," *CCF Transactions on Pervasive Computing and Interaction*, vol. 4, no. 3, pp. 298–318, Sep. 2022, doi: 10.1007/s42486-022-00089-3.
[14] G. Xi, S. Lin, and D. Zou, "A hybrid model of regional path loss of wireless signals through the wall," *KSII Transactions on Internet and Information Systems*, vol. 16, no. 9, pp. 3194–3210, Sep. 2022, doi: 10.3837/tiis.2022.09.020.
[15] L. Polak, S. Rozum, M. Slanina, T. Bravenec, T. Fryza, and A. Pikrakis, "Received signal strength fingerprinting-based indoor location estimation employing machine learning," *Sensors*, vol. 21, no. 13, p. 4605, Jul. 2021, doi: 10.3390/s21134605.
[16] M. As, H. Shimizu, B. Benaissa, K. Yoshida, and M. Koppen, "Calibration Cost Reduction of Indoor Localization Using Bluetooth Low Energy Beacon," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 26, no. 1, pp. 97–106, Jan. 2022, doi: 10.20965/jaciii.2022.p0097.
[17] J. Bi *et al.*, "Fast radio map construction by using adaptive path loss model interpolation in large-scale building," *Sensors (Switzerland)*, vol. 19, no. 3, p. 712, Feb. 2019, doi: 10.3390/s19030712.
[18] I. Ashraf, S. Din, S. Hur, and Y. Park, "Wi-Fi positioning dataset with multiusers and multidevices considering spatio-temporal variations," *Computers, Materials and Continua*, vol. 70, no. 3, pp. 5213–5232, 2022, doi: 10.32604/cmc.2022.018707.
[19] B. Wang, X. Liu, B. Yu, R. Jia, and X. Gan, "An improved WiFi positioning method based on fingerprint clustering and signal weighted euclidean distance," *Sensors (Switzerland)*, vol. 19, no. 10, p. 2300, May 2019, doi: 10.3390/s19102300.
[20] J. Torres-Sospedra, R. Montoliu, S. Trilles, Ó. Belmonte, and J. Huerta, "Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems," *Expert Systems with Applications*, vol. 42, no. 23, pp. 9263–9278, Dec. 2015, doi: 10.1016/j.eswa.2015.08.013.
[21] W. Zhang, K. Liu, W. Zhang, Y. Zhang, and J. Gu, "Deep Neural Networks for wireless localization in indoor and outdoor environments," *Neurocomputing*, vol. 194, pp. 279–287, Jun. 2016, doi: 10.1016/j.neucom.2016.02.055.
[22] K. S. Kim, S. Lee, and K. Huang, "A scalable deep neural network architecture for multi-building and multi-floor indoor localization based on Wi-Fi fingerprinting," *Big Data Analytics*, vol. 3, no. 1, p. 4, Dec. 2018, doi: 10.1186/s41044-018-0031-2.
[23] J. W. Jang and S. N. Hong, "Indoor Localization with WiFi Fingerprinting Using Convolutional Neural Network," in *International Conference on Ubiquitous and Future Networks, ICUFN*, vol. 2018-July, pp. 753–758, Jul. 2018, doi: 10.1109/ICUFN.2018.8436598.
[24] D. Sun, E. Wei, L. Yang, and S. Xu, "Improving Fingerprint Indoor Localization Using Convolutional Neural Networks," *IEEE Access*, vol. 8, pp. 193396–193411, 2020, doi: 10.1109/ACCESS.2020.3033312.
[25] X. Song *et al.*, "Cnnloc: Deep-learning based indoor localization with wifi fingerprinting," *Proceedings - 2019 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Internet of People and Smart City Innovation, SmartWorld/UIC/ATC/SCALCOM/IOP/SCI 2019*, pp. 589–595, 2019, doi: 10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00139.
[26] X. Wang, L. Gao, S. Mao, and S. Pandey, "DeepFi: Deep learning for indoor fingerprinting using channel state information," in *2015 IEEE Wireless Communications and Networking Conference, WCNC 2015*, pp. 1666–1671, Mar. 2015, doi: 10.1109/WCNC.2015.7127718.
[27] H. Chen, Y. Zhang, W. Li, X. Tao, and P. Zhang, "ConFi: Convolutional Neural Networks Based Indoor Wi-Fi Localization Using Channel State Information," *IEEE Access*, vol. 5, pp. 18066–18074, 2017, doi: 10.1109/ACCESS.2017.2749516.
[28] A. B. Mazlan, Y. H. Ng, and C. K. Tan, "A Fast Indoor Positioning Using a Knowledge-Distilled Convolutional Neural Network (KD-CNN)," *IEEE Access*, vol. 10, pp. 65326–65338, 2022, doi: 10.1109/ACCESS.2022.3183113.

[29] A. B. Mazlan, Y. H. Ng, and C. K. Tan, "Teacher-Assistant Knowledge Distillation Based Indoor Positioning System," *Sustainability (Switzerland)*, vol. 14, no. 21, p. 14652, Nov. 2022, doi: 10.3390/su142114652.

[30] G. Hinton, O. Vinyals, and J. Dean, "Distilling the Knowledge in a Neural Network," *arXiv preprint arXiv:1503.02531*, 2015.

[31] A. N. Nor Hisham, Y. H. Ng, C. K. Tan, and D. Chieng, "Hybrid Wi-Fi and BLE Fingerprinting Dataset for Multi-Floor Indoor Environments with Different Layouts," *Data*, vol. 7, no. 11, p. 156, Nov. 2022, doi: 10.3390/data7110156.

[32] W. Son, J. Na, J. Choi, and W. Hwang, "Densely Guided Knowledge Distillation using Multiple Teacher Assistants," in *Proceedings of the IEEE International Conference on Computer Vision*, Oct. 2021, pp. 9375–9384, doi: 10.1109/ICCV48922.2021.00926.

BIOGRAPHIES OF AUTHORS

Aqilah Mazlan was born in 1999, in Perak, Malaysia. At present, she is an Electronics Engineering student at the Faculty of Engineering, Multimedia University, Malaysia. Her research interests are machine learning and indoor positioning. She can be contacted at email: aqilahmazlan99@gmail.com.

Yin Hoe Ng received the B.Eng. degree (Hons.) in Electronics Engineering, Master of Engineering Science and Ph.D. degrees from Multimedia University in 2004, 2008, and 2013, respectively. He is currently a senior lecturer with the Faculty of Engineering, Multimedia University. Besides that, he is also a Chartered Engineer (CEng) and Professional Technologist (P.Tech.) registered with Engineering Council United Kingdom and Malaysia Board of Technologists, respectively. His current research interests include advanced signal processing techniques for digital communication systems, machine learning, and indoor positioning. He can be contacted at email: yhng@mmu.edu.my.

Chee Keong Tan received the B.Eng. degree in Electronics (Telecommunication), M.Eng.Sc. degree in Information, Communication and Technology and the Ph.D. degree in Information, Communication and Technology from Multimedia University, Malaysia. He is currently a senior lecturer with the School of Information Technology, Monash University, Malaysia. He has carried out projects for telecommunication companies and cellular service providers, which led to the development of a few patents on wireless algorithm and protocol design. He is the main contributing authors to more than 20 international journal papers. His current research interests include radio resource management, 5G networks, indoor positioning scheme, game theory, machine learning, and artificial intelligence. He can be contacted at email: tan.cheekeong@monash.edu.