



Offset-aware transformer for continuous Wi-Fi fingerprinting localization

Min-Jae Kim¹ · Yoon-Sang Han² · Dong-Hoan Seo[†]

(Received December 8, 2025 ; Revised December 22, 2025 ; Accepted December 29, 2025)

Abstract: Wi-Fi fingerprint-based indoor localization is widely deployed due to its compatibility with existing wireless infrastructure. However, conventional fingerprinting is designed as a classification task over discretized reference points (RPs), making it difficult to infer continuous positions between RP grids. This structural limitation degrades both localization accuracy and spatial resolution, and retraining is required whenever new RPs are added, reducing scalability. In this paper, we propose Offset-Aware Transformer Localization (OTL), a new indoor localization framework that preserves the conventional RP grid-based structure while estimating continuous positions through offset regression. OTL employs a Transformer-based cross-attention mechanism, where measured RSS values are used as queries and radio map embeddings are used as keys and values, enabling the model to attend to spatially relevant RPs. The final position is predicted by regressing a continuous offset from the classified RP coordinate. Experiments conducted on the AI-Hub indoor localization dataset, collected from underground shopping malls and metro stations in Seoul, demonstrate that OTL improves localization accuracy compared to classification-based methods, achieving 38.02% accuracy and an average positioning error of 2.88 m. These results confirm that OTL effectively overcomes the inherent discretization constraints of fingerprinting and enables scalable localization without retraining when new RPs are added.

Keywords: Indoor positioning systems, Wi-Fi fingerprinting, Classification, Indoor location-based service, Deep Learning

1. Introduction

Modern society already utilizes various services based on users' location information. Indoor Location-Based Services (ILBS), which serve as the foundation of such services, play an important role in various indoor environments such as large shopping malls, subway stations, and logistics centers. The Global Positioning System (GPS), which estimates global positions based on satellites, has already become widespread. However, GPS performance degrades in indoor environments due to signal blockage and multipath effects, and therefore Wi-Fi-based alternative technologies are widely used for indoor localization [1]. In particular, Wi-Fi fingerprinting has low deployment cost and high service applicability because it can utilize the existing Wi-Fi infrastructure without additional facilities. This approach is considered based on universal Wi-Fi environments in facilities such as shopping centers, underground malls, and convention

centers. Based on this, ILBS research has been expanded from a focus on monitoring and control to analysis for operational perspectives.

Wi-Fi signals enable localization modeling because the signal characteristics change according to the distance from the access point (AP) to the receiving terminal. Thus, localization estimation can be performed based on mathematical modeling such as trilateration, including Time Difference of Arrival (TDOA) [2], Frequency Difference of Arrival (FDOA) [3], and Angle of Arrival (AOA) [4]. However, these methods have limitations in Non-Line of Sight (NLOS) environments due to Wi-Fi multipath fading [5]. In contrast, Wi-Fi fingerprinting has high reliability even in NLOS environments because it estimates positions by utilizing the patterns between the collected signals and the current signal, making it robust in any indoor environment.

Fundamentally, fingerprinting follows a localization-as-

[†] Corresponding Author (ORCID: <http://orcid.org/0000-0003-3610-0356>): Professor, Division of Electronics and Electrical Information Engineering & Interdisciplinary Major of Maritime AI Convergence, Korea Maritime & Ocean University, 727, Taejong-ro, Yeongdo-gu, Busan 49112, Korea, E-mail: dhseo@kmou.ac.kr, Tel: 051-410-4412

1 M. S. Candidate, Department of Electrical and Electronical Engineering & Interdisciplinary Major of Maritime AI Convergence, Korea Maritime & Ocean University, E-mail: kminjae2926@gmail.com, Tel: +82-51-410-4822

2 M. S., Department of Electrical and Electronical Engineering & Interdisciplinary Major of Maritime AI Convergence, Korea Maritime & Ocean University, E-mail: hanyasang@gmail.com, Tel: +82-51-410-4822

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

classification approach. To estimate the space between reference points that have not been measured, regression-based approaches using k-Nearest Neighbor (k-NN) have been adopted from early studies to the present [6][7]. However, due to the success of recent deep architectures, localization models have been developed and thus the performance of classification-based approaches has dramatically improved [8].

However, Wi-Fi fingerprinting performs location estimation in the form of classification based on Received Signal Strength Indication (RSSI) collected at Reference Points (RP). Therefore, despite the fact that the real-world environment is continuous, it has a structural limitation that only discontinuous position estimation on the RP grid is possible. This limitation requires dense RP deployment for high-resolution localization, which significantly increases deployment costs and operational burden.

Deep Learning (DL)-based models combine an encoder network for feature extraction and a classification network for location estimation. As in other fields, DL-based localization models have used various networks to extract good features from RSS. Recently, CNN [9]-[11], RNN [12]-[14], and Transformer-based structures [15]-[17] have been used to enhance RSS feature representation in indoor localization studies. Particularly, Transformer architectures are effective for learning correlations between signals. However, most studies still remain in the class-based RP selection framework, and thus do not solve the problem of continuous location estimation. As a result, localization models have low spatial representation capability and continue to rely on k-NN or regression approaches, similar to previous methods.

Therefore, in this paper, we propose the Offset-Aware Transformer Localization (OTL) model, which overcomes the limitations of classification-based models by estimating continuous locations and regressing fine-grained positions between RPs. First, the proposed model accurately estimates positions by learning effective RSS features through an embedding-localization branch. In this process, in parallel, the model applies the Transformer Cross-Attention mechanism to utilize the measured RSS as a Query and the radio map as Key/Value, focusing on RPs that are spatially relevant, and regresses offsets to the classified RP coordinates to output the final continuous coordinates. This model is designed in an end-to-end manner to improve integration among networks and is configured to utilize the existing radio map without modification. Through this process, our proposed model overcomes the rigidity of classification-based localization and improves fine-grained position representation,

enabling precise localization services. We prove improved performance by comparing the proposed model with other models.

The main contributions of this article are as follows:

- We propose an OTL model based on Wi-Fi RSS fingerprinting that maintains the accuracy of classification models while expressing fine-grained positions.
- We design a Cross-Attention-based RP selection method that uses RSS as a Query and the radio map as Key/Value to reflect spatial relevance.
- Through experiments using a public AI-Hub dataset, we demonstrate improved performance compared to existing classification-based models, achieving 38.02% accuracy and an average error of 2.88 m.

2. Related Work and Limitations

Wi-Fi fingerprint-based indoor localization estimates user positions by exploiting the similarity patterns of RSSs. In general, the distance between the measured RSS and a preconstructed radio map is computed, and the location is estimated using kNN [8][18]. However, since fingerprinting represents locations only at RP grid levels, it inherently suffers from limited spatial resolution and cannot accurately estimate positions in the regions between RPs.

Recently, numerous studies [9]-[17][26] have employed deep learning architectures composed of feature extractors and classifiers for localization. CNN-based approaches [9]-[11] and RNN-based approaches [12]-[14] significantly improved localization accuracy by learning the nonlinear characteristics of RSS measurements. Nevertheless, these methods remain constrained to RP-based classification, resulting in limited spatial expressiveness. Improving spatial resolution requires finer RP grids, which in turn leads to an exponential increase in the cost of radio map construction. Transformer-based models [15]-[17][27] have shown potential for enhanced spatial representation by modeling correlations among signals; however, most of these approaches still rely on RP classification results as final location estimates, leaving the problem of continuous localization unresolved.

To mitigate these limitations, kNN-based localization methods [20]-[26] encode RSS measurements using neural networks and perform localization via kNN search in the learned feature space. This approach effectively leverages the representational power of deep learning while alleviating strict dependence on RP grids. Alternatively, regression-based methods that directly infer coordinates have also been proposed [28]-[32]. Since these methods

are generally not end-to-end, they require distance-based loss functions for training, and deep metric learning has become a common practice in fingerprinting-based localization. However, such approaches do not fully exploit the strong classification capability of deep learning, and localization accuracy remains limited. Therefore, for practical IPS deployment, a model capable of precise localization in continuous space—similar to KNN—while effectively utilizing deep learning’s classification power is required.

3. Proposed Method

3.1 Overview of the Proposed Model

Fingerprinting requires collecting RSS during the deployment phase and processing it into a radio map. Since reference points (RPs) form the basis of localization accuracy, they must be carefully selected. Although physical space is continuous, digital systems inevitably represent it discretely; therefore, how the RPs are defined directly determines the achievable resolution of the service. kNN-based similarity estimation can regress intermediate positions between RPs, but its precision is limited due to noise and multipath fading in Wi-Fi signals. In contrast, DL-based classification approaches can achieve high accuracy but cannot estimate fine-grained positions due to structural limitations. To address this issue, the proposed model aims to enhance spatial resolution by regressing offsets around the classified RP, rather than increasing the density of RPs. This presents a new challenge in overcoming the inherent limitations of classification-based localization. Therefore, we propose Offset-Aware Transformer Localization (OTL), which predicts continuous positions between discrete RPs.

The proposed OTL model consists of two parallel network branches. The first is a fingerprint branch that extracts features from online RSS and estimates location through RP classification. The second is an offset branch that uses the same RSS input to estimate the distance offset from the classified RP. The overall architecture is illustrated in **Figure 1**. The outputs of the two branches are integrated to produce the final location estimate.

The fingerprinting process includes an offline phase where RSS is collected and the model is trained. During training, a center module learns optimized radio map representations for each RP. Therefore, the proposed model does not require additional radio map alignment or reconstruction. Moreover, the learned embedding structure is directly used during the online phase, minimizing the computational load associated with radio map

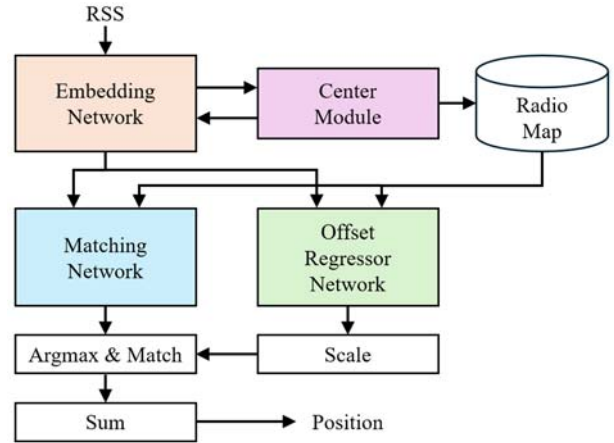


Figure 1: Architecture of the proposed deep similarity model

size. By integrating the classification and offset results, the user’s location is predicted more precisely. The two branches operate in parallel and are trained in an end-to-end manner, improving training efficiency and model consistency.

3.2 Fingerprint Classification Branch

According to the assumption of fingerprinting, the location is estimated by comparing RSS from APs with radio map data. The data distance has relatively low accuracy due to mismatches between signal distance and actual path distance caused by multipath fading. Therefore, as mentioned in Section 2, existing studies have improved accuracy by embedding RSS features to re-range signals and normalize distance measures. The proposed model extends the localization structure of Lee *et al.* [27].

The embedding network of the model takes an input dimension equal to the number of APs and utilizes a wider embedding layer to extract richer features. RSS is the result of Wi-Fi scanning measured by the device, and is defined as

$$\mathbf{x} \in \mathbb{R}^{N_{AP}}. \quad (1)$$

In general, a radio map R refers to the collection of RSS values at RPs, but the proposed model infers it using the Center module. Therefore, the radio map has the same embedding dimension N_d :

$$R = \{r_1, r_2, \dots, r_{N_{RP}}\}, \quad (2)$$

where $\mathbf{r}_i \in \mathbb{R}^{N_d}$ is the embedded RSS of the i -th RP and N_{RP} denotes the number of reference points. The radio map learned during the offline training phase is stored in a separate database, and in the online phase, inference is performed using only this stored data without the Center module.

To normalize RSS depending on device and service environment characteristics, we apply min-max scaling as follows

$$x_{b,j} = \frac{RSS_{b,j} - \sigma_{Min}}{\max(RSS_j) - \sigma_{Min}}, \quad (3)$$

where b and j denote batch and AP indices, respectively. Wi-Fi IoT devices typically do not receive signals weaker than -95 dBm, so we set $\sigma_{Min} = -96$ dBm. The maximum RSS value $\max(RSS_j)$ is measured based on the environment and device. The normalized RSS has a shape $x \in \mathbb{R}^{B \times N_{AP}}$ and we assume $B = 1$ in this study.

The embedding network is composed of fully connected layers. As the number of APs increases, the node width must also increase because the feature space becomes larger. Since RSS has a strong correlation with distance according to the Friis equation, a shallow network is sufficient; thus, a 3-layer structure is used. The embedded feature vector $X \in \mathbb{R}^{B \times N_d}$ has a hidden node dimension of N_d . N_d varies depending on the number of APs and model complexity. All blocks of the embedding network have equal width except for the input layer. Each layer output X_l is defined as

$$X_l = \text{ReLU}(W_l X_{l-1}^T + b_l), \quad (4)$$

where W_l and b_l represent learnable weights and biases. The overall embedding is expressed as

$$X = \text{Embedding}(x). \quad (5)$$

For the matching network of the localization branch, RSS and radio map embeddings are used as inputs. As shown in **Figure 1**, the matching network computes similarity using an attention mechanism. The two embedding vectors are normalized using

$$\text{Norm}(X) = \frac{X - \text{Embedding}(x)}{\sqrt{\text{Var}(X) + \epsilon}} \times \gamma + \beta, \quad (6)$$

where γ and β are learnable affine parameters, and ϵ is a small constant for numerical stability. Layer normalization is applied. The similarity matrix is computed as

$$S = \text{softmax}\left(\frac{\text{Norm}(X_E) \text{Norm}(X_R^T)}{\sqrt{d_x}}\right), \quad (7)$$

where d_x is the embedding dimension. The predicted location index is determined as

$$l = \text{argmax}(S). \quad (8)$$

The localization branch uses the similarity vector S as the learning result.

3.3 Offset Regression Branch

In the proposed OTL model, the classification branch outputs the predetermined RP location, and thus it can only estimate discrete RP-grid-based positions. However, the actual indoor space is continuous, and accurate estimation of a user's location between RPs requires fine error correction. To address this, the proposed model introduces a parallel offset regression branch that regresses continuous spatial coordinates around the classified RP.

Figure 2 illustrates the proposed regression network.

The offset branch is based on a Transformer structure, where the measured RSS is used as the Query and the radio map embedding is used as the Key/Value to focus on spatially related RPs. Similar to the localization branch, the RSS acts as the Query while the radio map serves as the Key/Value. Using a Cross-Attention mechanism, Query, Key, and Value are defined as

$$Q = W_Q X_E, \quad (8)$$

$$K = W_K X_R \text{ and} \quad (9)$$

$$V = W_V X_R, \quad (10)$$

where X_E is the online RSS embedding, X_R is the radio map embedding, and W_Q, W_K, W_V are learnable matrices. Cross-attention weights are computed as

$$A = \text{softmax}\left(\frac{QK^T}{\sqrt{d_x}}\right), \quad (11)$$

where A represents the attention distribution over RPs. The attention output is computed as

$$H = AV. \quad (12)$$

The feature H is passed to the Offset Head, composed of fully connected layers, to regress the offset from the classified RP coordinate.

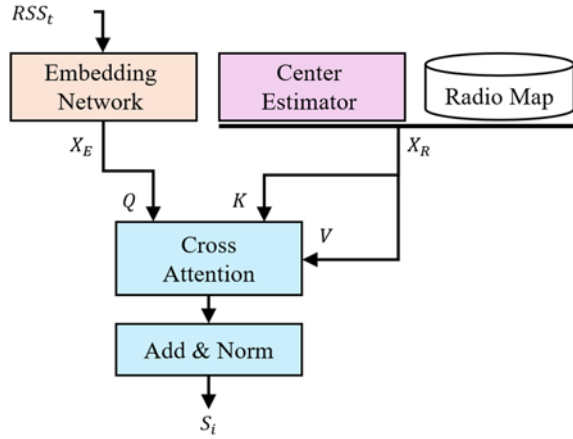


Figure 2: Structure of the proposed Offset regression branch

$$\Delta p = W_o H + b_o. \quad (13)$$

Finally, the continuous final position is obtained as

$$\hat{p} = p_\ell + \Delta p. \quad (14)$$

The Offset Regression Branch complements the structural limitations of classification-based localization and enables continuous spatial position representation between RPs.

3.4 Loss & Joint Optimization

The proposed OTL model jointly learns a classification branch and an offset regression branch. The classification branch determines the RP location, while the offset regression branch refines the location into continuous coordinates around the classified RP. The two branches operate in parallel and the entire network is trained in an end-to-end manner.

Additionally, the network employs Center Loss to compact the RSS embedding around the center of each RP cluster. The Center Loss \mathcal{L}_{cnt} is defined based on the center embedding $\mathbf{X}_c \in \mathbb{R}^{B \times I}$ as

$$\mathcal{L}_{cnt} = \frac{1}{2B} \sum_{b=1}^B \|X_b - X_c\|_2^2. \quad (15)$$

The classification branch is trained using the cross-entropy loss

$$\mathcal{L}_{cls} = - \sum_{i=1}^{N_{RP}} y_i \log(\hat{c}_i). \quad (16)$$

The offset regression branch is trained using an L1 distance loss

$$\mathcal{L}_{reg} = \|p - \hat{p}\|_1. \quad (17)$$

Finally, the overall optimization objective is formulated as

$$\mathcal{L} = \mathcal{L}_{cls} + \lambda_1 \mathcal{L}_{reg} + \lambda_2 \mathcal{L}_{cnt}, \quad (18)$$

where λ_1 and λ_2 balance the contributions of each loss term. We set λ_1 and λ_2 to 1 and 0.5, respectively. Although the classification loss and regression loss exhibit different convergence behaviors, they operate on similar scales, making it reasonable to assign comparable weights. In contrast, the center loss is introduced to stabilize the embedding distribution, and an excessively large weight can adversely affect the localization learning process. Therefore, a relatively small weight is preferred for the center loss. This design maintains high classification accuracy while improving continuous spatial representation between RPs and stabilizing the RSS embedding distribution.

4. Database Experiment

4.1 Environment and Implementation Detail

In this paper, we use the ‘‘Indoor Localization Fusion Dataset Construction’’ provided by AI-Hub. This dataset consists of Wi-Fi signal measurements collected in complex indoor environments such as underground shopping malls and subway stations in Seoul, reflecting realistic commercial deployment conditions. Each RP contains repeatedly measured RSS values from multiple APs. We select 10 locations from the dataset for our experiments. **Table 1** summarizes the detail of dataset for each region.

The selected environments include underground corridors, open spaces with pillars, and irregular single-floor layouts, while other characteristics vary across sites. On average, the dataset includes 997 RPs and 138 APs per site.

Table 1: Dataset Description For Each Region

Region Name	Area Size (m ²)	Number of AP	Number of RP
Euljiipgu	2,271	142	531
Jonggak	4,724	276	1,149
Jongo	4,165	115	1,193
Majeongyo	1,210	38	178
Dongdaemun	2,151	111	423
Myeongdong	3,838	27	850
Myeongdong Stn.	2,215	112	561
Yeongdeunpo Stn.	4,678	125	1,347

We utilize the 2D coordinates of each RP and the corresponding RSS values for training and evaluation. RSS values are highly influenced by AP placement, device orientation, obstacles, and signal path variations. Multipath fading and NLOS effects are particularly strong in underground spaces. These characteristics make the dataset well-suited for evaluating Wi-Fi-based indoor localization models and validating the generalization performance of the proposed method.

In fingerprinting-based localization, data collection must be performed for all RPs; therefore, the dataset is split at the RSS sample level. In this study, a total of 250 RSS samples were divided into training and validation sets with a ratio of 200:50. The tuning process of the proposed model is performed on indoor mall data from central Seoul. The training set is used for parameter learning, while the validation set monitors convergence and adjusts hyperparameters. The test set is used for final performance evaluation.

The proposed OTL model is implemented using PyTorch and all experiments are conducted on an NVIDIA GPU platform. Before being fed into the embedding network, RSS values are normalized using min-max scaling to mitigate device-dependent variations. The measurements are collected from public Wi-Fi APs operating in the 2.4 GHz band based on IEEE 802.11ac or higher. The minimum signal sensitivity is set to -96 dBm, and maximum RSS values are derived from the observed distributions of each AP.

The embedding network consists of three Fully Connected Layers, and the hidden dimension N_{dis} selected based on both the number of APs and the required model capacity. The offset regression branch adopts a Transformer-based Cross-Attention structure using the Query–Key–Value mechanism. Training is conducted using the Adam optimizer with an empirically chosen initial learning rate. The model is trained end-to-end using a combination of Center Loss, Classification Loss, and Offset Regression Loss.

Unless otherwise stated in ablation studies, the network depth and width are set to 4 and 800, respectively. While the same architecture is applied to all sites for fair comparison, the optimal network configuration may differ depending on environmental characteristics.

4.2 Performance Evaluation of the Proposed Model

In this section, we evaluate the performance of the proposed OTL model by comparing it with kNN-based approaches [19, 1], an RNN-based learning approach [12], and metric-learning-

based methods [23, 27]. Two evaluation metrics are employed: classification accuracy of RP labels and the average positioning error calculated using the Euclidean distance. The kNN-based algorithms utilize Euclidean distance [1] and cosine similarity [19]. MIMO [12] considers localization as a classification problem by mapping RSS to RP classes. DeepMetricFi [23] minimizes the discrepancy between signal and spatial embeddings and performs WKNN for location estimation. Additionally, DSN [27] integrates metric learning with a deep classification model.

Figure 3 illustrates the cumulative distribution function (CDF) of localization errors. For fair comparison, the number of neighbors K is fixed to 9 across all methods. Traditional distance-based approaches (Euclidean and Cosine) show similar performance, while MIMO performs better owing to its classification capability, although the improvement margin remains limited. DeepMetricFi improves performance by adding triplet loss and path distance constraints but is still restricted by its dependence on WKNN. DSN significantly enhances accuracy through a metric-learning-enabled encoder followed by a classifier.

A deeper investigation reveals that kNN-based and MIMO approaches tend to yield lower and evenly distributed performance, whereas the top three methods—DeepMetricFi, DSN, and the proposed OTL—demonstrate consistently superior results. Among these, DeepMetricFi exhibits relatively higher accuracy in short-distance error ranges; however, overall accuracy remains lower due to regression constraints. The proposed model refines final localization results using a regression branch, achieving less than 7 m error in approximately 90% of test cases, outperforming all other methods.

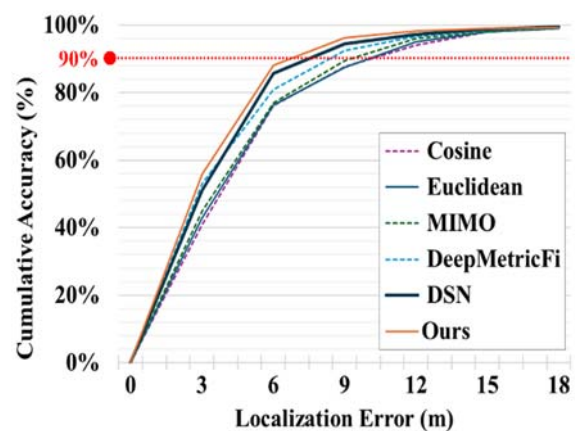


Figure 3: Cumulative distribution functions of our method and the comparisons

Table 2: Performance Accuracy And Error Obtained Via Our Method And Comparisons

Method	Localization Accuracy (%)	Localization Error Distance (m)
Cosine [19]	20.74	3.80
Euclidean [1]	19.33	3.81
MIMO [12]	20.18	3.49
DeepMetricFi [23]	29.17	3.25
DSN [27]	30.03	3.01
Ours	30.02	2.88

Table 2 summarizes the localization accuracy and average error for each method. Classification accuracy indicates the percentage of correctly predicted RP labels. Results reveal discrepancies between accuracy and distance errors in some models: classification-based models tend to mispredict distant locations when errors arise, whereas metric learning-based models reduce spatial deviation despite lower accuracy. The proposed model achieves 30.02% accuracy and an average positioning error of 2.88 m, demonstrating a superior balance between both metrics. Although DSN uses the same classifier as ours, the proposed OTL achieves better distance performance owing to its final offset regression refinement stage.

4. Conclusion

In this paper, we proposed a novel Offset-Aware Transformer Localization (OTL) model to overcome the inherent limitation of Wi-Fi RSS fingerprinting, where localization is restricted to discrete RP grids. The proposed architecture integrates a classification branch for estimating RP locations and a parallel regression branch employing cross-attention to infer continuous fine-grained positions between RPs. In addition, center loss is applied to enforce compact feature embedding and improve both classification and regression stability. Experimental results demonstrate that OTL achieves superior and well-balanced performance in both classification accuracy and average localization error compared with conventional distance-based and deep learning-based approaches.

Despite these advantages, several limitations remain. First, Wi-Fi RSS data are highly sensitive to environmental changes, AP deployment modifications, and temporal fluctuations, which may degrade model performance. Second, while the regression

branch enhances spatial continuity, it still struggles to fully capture complex environmental factors that influence radio propagation. Lastly, experiments were conducted on datasets from limited indoor regions; hence, further verification is necessary to confirm scalability in large-scale commercial spaces.

Future work will address these limitations by incorporating additional signals such as BLE, UWB, and geomagnetic sensing to construct a more robust multimodal indoor localization framework. Furthermore, we will adopt online learning and environment-adaptive calibration techniques to maintain localization stability under environmental variations and AP reconfigurations. We also plan to enhance continuous localization performance by extending Transformer capabilities, including improved self-attention mechanisms and radio-map refinement strategies. Finally, we will validate the generalization and practical deployment feasibility of the proposed model in extensive and diverse real-world indoor environments.

In conclusion, the proposed OTL model successfully overcomes the structural limitations of classification-based Wi-Fi localization systems, enabling continuous spatial representation between RPs, and is expected to contribute to improved precision in next-generation Indoor Location-Based Services (ILBS).

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00352187) and Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE) (No. RS-2024-00424595).

Author Contributions

Conceptualization, M. J. Kim and Y. S. Han; Methodology, M. J. Kim; Software, Y. S. Han; Validation, M. J. Kim and Y. S. Han; Formal Analysis, M. J. Kim; Data Curation, Y. S. Han; Writing—Original Draft Preparation, M. J. Kim; Writing—Review & Editing, Y. S. Han; Visualization, D. H. Seo; Supervision, Corresponding Author; Project Administration, D. H. Seo.

References

- [1] J. H. Seong, E. C. Choi, J. S. Lee, and D. H. Seo, "High-speed positioning and automatic updating technique using Wi-Fi and UWB in a ship," *Wireless Personal Communications*, vol. 94, no. 3, pp. 1105-1121, 2017.

- [2] S. G. Nagarajan, P. Zhang, and I. Nevat, "Geo-spatial location estimation for Internet of Things (IoT) networks with one-way time-of-arrival via stochastic censoring," *IEEE Internet Things Journal*, vol. 4, no. 1, pp. 205-214, 2017.
- [3] K. Lin, W. Wang, Y. Bi, M. Qiu, and M. M. Hassan, "Human localization based on inertial sensors and fingerprints in the Industrial Internet of Things," *Computer Networks*, vol. 101, pp. 113-126, 2016.
- [4] 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, 2018.
- [5] W. Watson and T. McElwain, "4D CAF for localization of co-located, moving, and RF coincident emitters," in *MILCOM 2016-2016 IEEE Military Communications Conference*, pp. 948-951, 2016.
- [6] D. Li, B. Zhang, and C. Li, "A feature-scaling-based k -nearest neighbor algorithm for indoor positioning systems," *IEEE Internet of Things Journal*, vol. 3, no. 4, pp. 590-597, 2016.
- [7] P. Chen, F. Liu, S. Gao, P. Li, X. Yang, and Q. Niu, "Smartphone-based indoor fingerprinting localization using channel state information," *IEEE Access*, vol. 7, pp. 180609-180619, 2019.
- [8] J. H. Seong and D. H. Seo, "Selective unsupervised learning-based Wi-Fi fingerprint system using autoencoder and GAN," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1898-1909, 2020.
- [9] J. W. Jang and S. N. Hong, "Indoor localization with wifi fingerprinting using convolutional neural network," in *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 753-758, 2018.
- [10] R. S. Sinha and S. H. Hwang, "Comparison of CNN applications for RSSI-based fingerprint indoor localization," *Electronics*, vol. 8, no. 9, p. 989, 2019.
- [11] Q. Ye, X. Fan, G. Fang, H. Bie, X. Song, and R. Shankaran, "CapsLoc: a robust indoor localization system with WiFi fingerprinting using capsule networks," in *ICC 2020-2020 IEEE International Conference on Communications*, pp. 1-6, 2020.
- [12] M. T. Hoang, B. Yuen, X. Dong, T. Lu, R. Westendorp, and K. Reddy, "Recurrent neural networks for accurate RSSI indoor localization," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10639-10651, 2019.
- [13] H. Sun, X. Zhu, Y. Liu, and W. Liu, "WiFi based fingerprinting positioning based on Seq2seq Model," *Sensors*, vol. 20, no. 13, p. 3767, 2020.
- [14] Q. Ren, Y. Wang, S. Liu, and X. Lv, "FSTNet: Learning spatial-temporal correlations from fingerprints for indoor positioning," *Ad Hoc Networks*, vol. 149, 103244, 2023.
- [15] Z. Zhang, H. Du, S. Choi, and S. H. Cho, "TIPS: Transformer based indoor positioning system using both CSI and DoA of WiFi signal," *IEEE Access*, vol. 10, pp. 111363-111376, 2022.
- [16] A. Jagannath and J. Jagannath, "Embedding-assisted attentional deep learning for real-world RF fingerprint of bluetooth," *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 4, pp. 940-949, 2023.
- [17] Z. Wu, P. Hu, S. Liu, and T. Pang, "Attention mechanism and LSTM network for fingerprint-based indoor location system," *Sensors*, vol. 24, no. 5, 1398, 2024.
- [18] J. Hu, D. Liu, Z. Yan, and H. Liu, "Experimental Analysis on Weight K -nearest neighbor indoor fingerprint positioning," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 891-897, 2019.
- [19] S. Bai, Y. Luo, M. Yan, and Q. Wan, "Distance metric learning for radio fingerprinting localization," *Expert Systems with Applications*, vol. 163, 113747, 2021.
- [20] A. Pandey, R. Sequeira, and S. Kumar, "SELE: RSS-based Siamese embedding location estimator for a dynamic IoT environment," *IEEE Internet of Things Journal*, vol. 9, no. 5, pp. 3672-3683, 2022.
- [21] L. Zhang, S. Wu, T. Zhang, and Q. Zhang, "Learning to locate: Adaptive fingerprint-based localization with few-shot relation learning in dynamic indoor environments," *IEEE Transactions on Wireless Communications*, vol. 22, no. 8, pp. 5253-5264, 2023.
- [22] J. Hu and C. Hu, "A WiFi indoor location tracking algorithm based on improved weighted k nearest neighbors and Kalman filter," *IEEE Access*, vol. 11, pp. 32907-32918, 2023.
- [23] P. Chen and S. Zhang, "DeepMetricFi: Improving Wi-Fi fingerprinting localization by deep metric learning," *IEEE Internet of Things Journal*, pp. 6961-6971, 2023.

- [24] S. Zhang, G. Zhang, R. Chen, and Y. Wang, "Multiple similarity analysis-based deep metric learning for enhancing Wi-Fi fingerprint indoor localization," *IEEE Internet of Things Journal*, vol. 11, no. 21, pp. 35681-35688, 2024.
- [25] A. K. Panja, S. F. Karim, S. Neogy, and C. Chowdhury, "Improving the sustainability of WiFi-enabled indoor localization systems through meta-heuristic based instance selection approach," *Expert Systems with Applications*, vol. 257, 125063, 2024.
- [26] S. H. Lee, W. Y. Kim, and D. H. Seo, "Automatic self-reconstruction model for radio map in Wi-Fi fingerprinting," *Expert Systems with Applications*, vol. 192, 116455, 2022.
- [27] S. H. Lee and D. H. Seo, "Scalable Wi-Fi fingerprinting localization by deep similarity network," *IEEE Internet of Things Journal*, vol. 12, no. 4, pp. 4197-4206, 2025.
- [28] L. Hao, B. Huang, H. Hong, B. Jia, and W. Li, "A channel adaptive WiFi indoor localization method based on deep learning," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1-6, 2021.
- [29] C. Y. Chen, A. I-Chi Lai, P. Y. Wu, and R. B. Wu, "Optimization and evaluation of multidetector deep neural network for high-accuracy Wi-Fi fingerprint positioning," *IEEE Internet of Things Journal*, vol. 9, no. 16, pp. 15204-15214, 2022.
- [30] A. Alitalashi, H. Jazayeriy, and J. Kazemitabar, "EA-CNN: A smart indoor 3D positioning scheme based on Wi-Fi fingerprinting and deep learning," *Engineering Applications of Artificial Intelligence*, vol. 117, 105509, 2023.
- [31] A. Brunello, A. Montanari, and N. Saccomanno, "Towards interpretability in fingerprint based indoor positioning: May attention be with us," *Expert Systems with Applications*, vol. 231, 120679, 2023.
- [32] S. L. Ayinla, A. A. Aziz, and M. Drieberg, "SALLoc: An accurate target localization in Wifi-enabled indoor environments via SAE-ALSTM," *IEEE Access*, vol. 12, pp. 19694-19710, 2024.