IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 6, June, 2026, pp. 1971-1977
Journal homepage: http://iieta.org/journals/isi

A Dual-Modality Deep Learning Framework with Received Signal Strength Indicator–Fine Timing Measurement Fusion for Accurate Wi-Fi Indoor Localization

Kahlaa K. Al-Nassrawy[1*], Ghaidaa A Al-Sultany[2]

[1] Department of Information Networks, College of Information Technology, University of Babylon, Babylon 51002, Iraq
[2] College of Engineering and IT, Al-Zahraa University for Women, Karbala 56001, Iraq

Corresponding Author Email: k.alnassrawy@gmail.com

Copyright: ©2026 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/isi.310619

Received: 16 March 2026
Revised: 29 May 2026
Accepted: 10 June 2026
Available online: 30 June 2026

Keywords:
indoor localization, Wi-Fi positioning, Received Signal Strength Indicator–Fine Timing Measurement fusion, deep learning, multimodal sensor fusion, Fine Timing Measurement, wireless sensing

ABSTRACT

Accurate indoor localization remains challenging due to signal attenuation, multipath propagation, and environmental variations that significantly affect wireless measurements. Wi-Fi-based Received Signal Strength Indicator (RSSI) fingerprinting provides a low-cost localization solution but suffers from signal instability and ambiguity in complex indoor environments. Fine Timing Measurement (FTM) offers complementary ranging information; however, its standalone performance is also influenced by noise and environmental conditions. This study proposes a dual-modality deep learning framework that integrates RSSI and FTM measurements for Wi-Fi indoor localization. The proposed framework adopts independent normalization for heterogeneous signal streams, followed by feature-level fusion and a dense–convolutional learning architecture to capture global dependencies and local signal correlations. The model was evaluated using the ESP32C3-RSSI-FTM dataset collected from indoor environments with multiple access points. Experimental results show that the proposed framework achieved a mean absolute error (MAE) of 0.0297 m, a root mean squared error (RMSE) of 0.0456 m, and a cosine similarity of 0.977 between estimated and reference positions. These results indicate that combining complementary RSSI and FTM information through a unified deep learning architecture can improve localization robustness compared with relying on a single measurement source. Although further validation across diverse buildings and deployment scenarios is required, the proposed framework provides a practical approach for intelligent indoor positioning systems based on low-cost wireless devices.

1. INTRODUCTION

The rapid development of indoor environments that include intelligent buildings, warehouses, and healthcare is increasing the need to have accurate and reliable indoor positioning systems [1, 2]. Many applications that are context-dependent, such as indoor navigation, asset tracking, and patient monitoring simply cannot work without precise spatial data [1]. Unlike in outdoor environments, where the Global Positioning System (GPS) technology normally exhibits a satisfactory level of accuracy, indoor environments pose greater challenges, as signals are attenuated, multipath-propagated, and may be affected by walls, furniture, and human movement [3-6]. These challenges are further compounded by the common usage of Internet of Things (IoT) technologies [4]. Complex indoor environments consist of an increasing number of dense networks of interconnected devices, which create large-scale heterogeneous wireless data streams [7]. As a result, indoor localization systems are expected to be not only accurate, but also scalable and robust in noisy and dynamically changing signal conditions [7]. Standard positioning approaches developed for idealized environments remain insufficient to address these specifications [7]. Classical methods such as triangulation and trilateration determine target location by measuring distance or angle from anchor nodes. While these approaches are effective in line-of-sight conditions, their performance degrades substantially in complex indoor settings. The reliability of these techniques is therefore limited by environmental constraints and measurement uncertainty, making them unsuitable for many real-world indoor deployments [7, 8]. Conversely, Wi-Fi-based localization has become a practical choice because it is cost-effective and reliant on existing wireless local area network (WLAN) infrastructure [9]. Among Wi-Fi-based approaches, fingerprinting has gained considerable attention because it does not require explicit distance estimation. Instead, it exploits distinctive patterns in wireless signal measurements—most commonly Received Signal Strength Indicator (RSSI) values—collected from multiple access points [9-11]. This localization framework typically involves two phases [10-12]. In the offline phase, RSSI readings are recorded at predetermined locations to construct a spatial radio map [10-12]. During the online phase, real-time RSSI measurements are matched against this radio map to determine the user's exact location [10-12]. Although it is practical, RSSI-based fingerprinting has a number of shortcomings. RSSI is very sensitive to changes in the environment, and in crowded indoor environments, it tends to be ambiguous [11]. New developments in wireless standards have offered Fine Timing Measurement (FTM), which provides better ranging information, using signal propagation time [13]. Although FTM provides better distance estimation, it is highly susceptible to noise and environmental variability when it is used alone [14]. These observations highlight the necessity of localization models that are capable of incorporating heterogeneous wireless measurements and at the same time are resilient to indoor noise and signal distortion. Instead of using the properties of one modality or just geometric considerations, the combination of complementary signal properties through learning-based approaches is a promising direction. Following this trend, the current study combines two complementary and inexpensive Wi-Fi measurements in a single learning model, which can be used in a commodity device with no extra infrastructure to enable an accurate indoor localization system.

The principal contributions presented in this paper are summarized as follows:

•Proposing a robust hybrid deep learning model combining RSSI and FTM signals from ESP32C3 devices to significantly enhance indoor positioning accuracy.

•Introducing an integrated lightweight architecture, which consists of an independent per-modality normalization followed by a hybrid dense–convolutional design, where the two heterogeneous signals are first balanced in scale and then fused by layers that learn their global and local dependencies, providing a simpler alternative to fusion methods based on separate correlation analysis.

•Developing a hierarchical neural network that leverages the fused RSSI and FTM data to capture the interdependency between global and local features through dense and convolutional layers to precisely localize in 2D indoor environments.

The remainder of this paper is organized as follows: A comprehensive review of existing work in the field of indoor localization and learning-based fingerprinting is provided in Section 2. The dataset description process and the construction of the dual-modality dataset are presented in Section 3. Section 4 presents the results and discussion. Section 5 presents the conclusion and future directions of the proposed model.

2. RELATED WORK

Indoor localization systems are very essential for smart environments and IoT applications, and have been widely studied recently. However, it remains difficult to obtain stable positioning in the indoor environment because of multipath propagation, signal attenuation, and the dynamics of the environment. To meet these challenges, numerous approaches have been explored. Existing literature can be categorized into three methodological types to better frame the current context.

Traditional Machine Learning Approaches: Recent research has shown that conventional machine learning algorithms such as K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and decision trees can be used to train decision boundaries with RSSI fingerprints and achieve superior localization performance compared to probabilistic models [15]. These methods, however, treat each fingerprint as a single static sample, and they do not consider the nonlinear relationships between features or complex signal structures. This imposes an inherent constraint on their resilience and usefulness in large or cluttered indoor environments [15].

Deep Learning and Feature Optimization: Deep learning approaches are gaining popularity to address the drawbacks of the conventional algorithms. Convolutional Neural Networks (CNNs) have been widely adopted to conduct indoor localization by transforming RSSI fingerprints into structural representations, and then extracting spatial features without manually engineering features [16]. In addition, another line of research in this category is the optimization of features and model efficiency. The most informative access points and the reduction of computational complexity have been accomplished using methods like dimensionality reduction and feature selection. For example, optimizing learning models has been demonstrated to improve the accuracy and stability of the input features before localization [17]. However, all these single-signal approaches—either with conventional CNN or customized CNN—are highly sensitive to the quality of the starting fingerprint representation.

Hybrid Architectures and Signal Fusion: Apart from single-signal learning, recent studies have been conducted on hybrid architectures. Some architectures integrate signal processing with neural networks to improve the robustness in noisy environments [18]. While these achieve significant improvements, they generally concentrate on one type of signal. In recent years, comparative studies based on surveys have pointed out the need for multimodal localization systems that are able to merge complementary wireless measurements [19]. These works highlight that while some spatial information can be obtained from measurements such as RSSI, CSI, and timing-based features, none of them are robust enough to be used individually. Most current schemes either take a single mode of measurement, or take fused features as inputs that do not model internal interactions. This paper's framework overcomes these problems by integrating RSSI and FTM data in a systematic way in an integrated computational architecture.

3. MATERIALS AND METHODS

This section presents the dataset used in this paper, the studied problem formulation, and the proposed model for robust indoor localization based on RSSI and FTM readings.

3.1 Dataset description

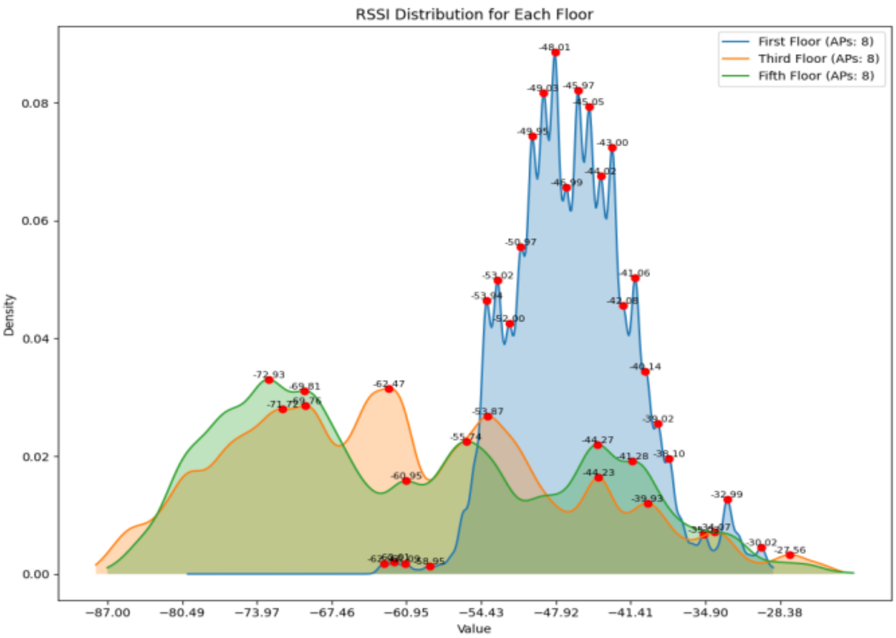
The dataset utilized in this paper is the ESP32C3 Wi-Fi FTM RSSI dataset [20], which was acquired from ESP32C3 devices (Wi-Fi enabled). These devices—in addition to measuring RSSI—incorporate the capability to extract FTM readings. The dataset environment consists of eight fixed devices acting as access points (APs) to collect synchronized RSSI and FTM readings from a mobile device moved by an autonomous robot Turtlebot3, to ensure high-precision ground truth while simulating realistic user trajectories. Readings are samples associated with a reference position (x, y) that serve as multimodal position labels to support AI techniques for localization tasks. The smartphone was mounted at a height of 75 cm to mimic the typical elevation of a device in a user's pocket. The dataset features were collected in a setting of a multi-floor building, covering diverse indoor layouts to collect the dataset entries. Table 1 illustrates basic statistics of the collected features.

Table 1. Features statistics for all floors

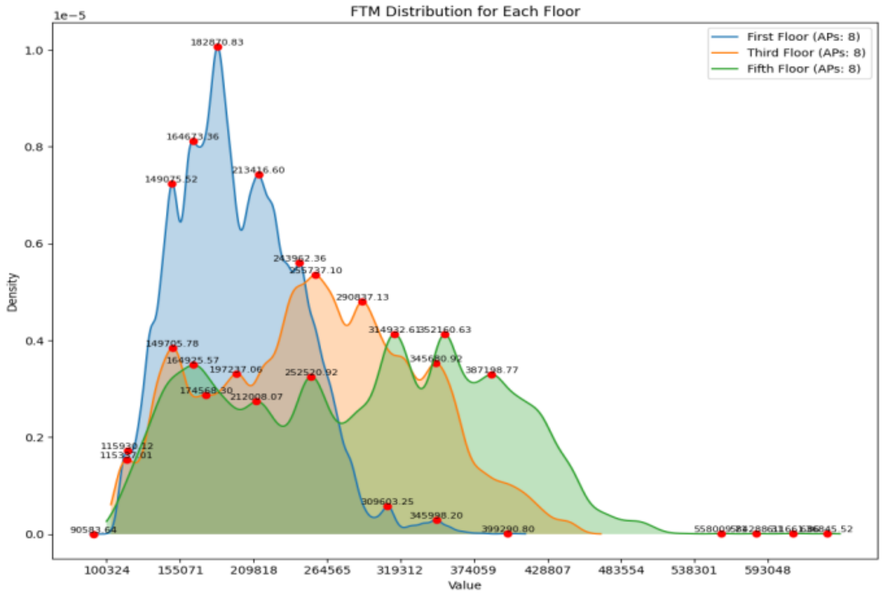
Floor	Mean	Median	Std	Min	Max
Received Signal Strength Indicator (RSSI)					
First Floor	-46.39	-47.0	5.32	-80.0	-29.0
Third Floor	-61.29	-63.0	13.67	-88.0	-23.0
Fifth Floor	-60.88	-64.0	14.51	-87.0	-22.0
Fine Timing Measurement (FTM)					
First Floor	197553.47	192339.0	44624.14	87984.0	412289.0
Third Floor	256475.96	257262.5	76205.14	103637.0	468531.0
Fifth Floor	287719.09	296708.5	97123.94	100324.0	646700.0

The values in Figure 1 show a non-homogeneous distribution of values, which indicates difficulty in combining those values to assemble a larger dataset for training.

Although the values of RSSI and FTM have common values areas, those common readings do not reflect the same spatial locations. Consequently, the same APs readings at an instance of time do not indicate the same location at each floor. The modality represented in this dataset includes RSSI and FTM. RSSI readings resemble the strength of the signal received by the user's equipment. The readings are recorded in decibel-milliwatts (dBm) and with ranges indicated in Table 1. On the other hand, FTM readings represent the round-trip timed distance features in meters. The benefit of FTM observations is that they act as assistant features to enhance localization due to their resilience against noise and other physical phenomena like multipath loss. In this dataset each row of readings (RSSI and FTM vectors) indicates a spatial 2D location (x, y).



(a) Received signal strength indicator (RSSI)



(b) Fine timing measurement (FTM)

Figure 1. Received signal strength indicator (RSSI) and fine timing measurement (FTM) values distribution

3.2 Problem formulation

The presented indoor localization task can be classified as a supervised regression problem. The goal is to estimate the localization of a user in a heterogeneous wireless environment. The dataset encoded two streams of readings from each access point, namely RSSI (r) and FTM (f) signals, where $r \in \mathbb{R}^{N_r}$ and $f \in \mathbb{R}^{N_f}$ represent RSSI and FTM streams respectively, while N_r and N_f indicate the number of access points in the current setting. The objective of the model is to approximate the nonlinear location mapping system.

$$\mathcal{F}: (r, f) \rightarrow l, where\ l = (x, y) \in \mathbb{R}^2 \quad (1)$$

The system is designed to optimize location accuracy by lowering the mean squared error (MSE) between the estimated and the ground truth locations provided in the dataset.

3.3 Data preprocessing

As seen in Table 1, the wireless signals under consideration, RSSI and FTM, have extremely different numerical ranges. This means that if these heterogeneous data streams are ingested into the network architecture directly without some sort of calibration, it creates huge scaling mismatches. An independent normalization approach is essential to create a common numerical domain and to ensure a fair representation of features. If this preprocessing is not performed, the optimization process may end up with a significant convergence bias, which is that the model is biased towards the spectrum of features that has large absolute values. Accordingly, each feature set was normalized independently using min–max scaling, which rescales the values of each modality into a common [0, 1] range. Min–max scaling was used instead of z-score standardization as it maintains the boundedness of the signals and both modalities would be scaled to the same range, which would result in a balanced contribution to the fusion and prevent the larger-scale modality from dominating. This separation is for balancing the training phase and also for ensuring a fair contribution from both signal sources. Also, manual feature engineering was not performed in this study, so that the model can learn spatial representations directly from the raw received signals.

3.4 Proposed hybrid localization model

The proposed model is a parallel-input neural network architecture that processes the RSSI and FTM signal streams simultaneously, making sure that the characteristics of each signal stream are preserved as shown in Figure 2. Each heterogeneous input stream is independently normalized using min–max scaling to avoid the dominance of the larger scale and to ensure a fair integration of features. After this key calibration, the normalized streams are then combined together to form a coherent and homogeneous feature map. This fused representation is then fed to a dense neural encoder with a particular focus on learning global interdependencies between the signals. The dense encoder is first applied because there is no natural topological order between access points in the concatenated RSSI–FTM vector and a convolution applied directly to raw fused values would not have any meaning. The dense layers thus embed the fused features into a learned representation that captures the global cross-modal dependencies and then unfold it into a structured sequence that the 1D convolution extracts local correlations from. The encoder uses a Leaky ReLU activation function (Eq. (2)) that can effectively prevent dead neurons.

$$f(x) = \max(ax, x) \quad (2)$$

where, a is a small non-negative hyperparameter to set the direction of the function's slope for negative input x. Two other techniques applied here: first, $L2$ regularization to improve generalization, and dropout to reduce overfitting. The resultant stream is fed into a 1D-CNN to learn localized dependencies across features. This convolution helps in assessing the exploitation of structured correlations among the fused features stream. The output regression layer acts as a spatial localization estimator for the input 2D location. The layout of the proposed model is illustrated in Figure 2, and Table 2 shows the parameters used in this model.

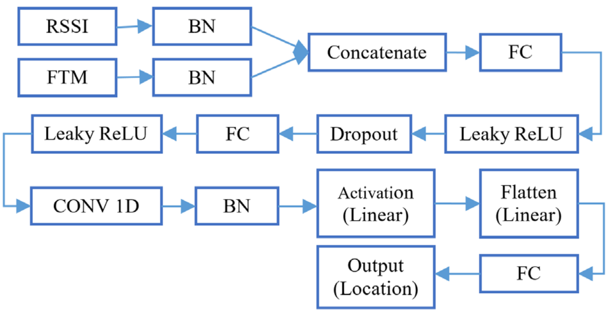


Figure 2. Dual-input RSSI–FTM localization model architecture

Note: RSSI = Received Signal Strength Indicator; FTM = Fine Timing Measurement.

Table 2. Structure and parameters of RSSI–FTM localization model

Layer	Type	Output Shape
RSSI Input	Input	(8,)
FTM Input	Input	(8,)
BN (RSSI)	BatchNorm	(8,)
BN (FTM)	BatchNorm	(8,)
Concatenation	Fusion	(16,)
Dense	FC + LeakyReLU	(128,)
Dropout	Regularization	(128,)
Dense	FC + LeakyReLU	(64,)
Reshape	–	(64, 1)
Conv1D	Kernel=3	(64, 32)
BN + ReLU	–	(64, 32)
Flatten	–	(2048,)
Dense	FC + ReLU	(32,)
Output	Linear	(2,)

Note: RSSI = Received Signal Strength Indicator; FTM = Fine Timing Measurement; BN = Batch Normalization; FC = Fully Connected; ReLU = Rectified Linear Unit.

3.5 Model training

For model training, the Adam optimizer was selected to update weights, while the MSE was employed as the loss function to guide the learning process and decrease the difference between true and estimated locations in Euclidean space as shown in Eq. (3).

$$MSE = \frac{1}{n}\sum_{1}^{n}\left(L_i - \hat{L}_i\right)^2 \qquad (3)$$

where, n is the batch size, L represents the true 2D location, and $\hat{L}$ is the 2D location estimated by the model.

For evaluating the model's predictive capability, the data was partitioned into training and unseen samples. Table 3 illustrates the hyper-parameters used to train the model. These training settings were selected to obtain a balance between stability and generalization. The Adam optimizer was chosen due to its adaptive learning behaviour that converged reliably on regression tasks with the common default learning rate of 0.001. Batch size is 32, which is a compromise between gradient stability and memory costs. The dropout rate is set to 0.2 and the L2 weight is 1×10^{-4}, which is a light regularization for the size of the dataset. Lastly, the maximum number of epochs is 750 and early stopping is used, which means that the training will automatically stop when the validation loss stops improving.

Table 3. Training parameters for the proposed model

Parameter	Value	Description
Optimizer	Adam	Adaptive Moment Estimation
Learning Rate	0.001	Fixed learning rate for all experiments
Loss Function	Mean squared error (MSE)	Regression loss for the coordinate estimation
Evaluation Metric	Mean absolute error (MAE)	Measures average localization error
Batch Size	32	Number of samples per training batch
Number of Epochs	750	Maximum training iterations
Early Stopping	Enabled	Prevents overfitting based on the validation loss
Dropout Rate	0.2	Applied after first dense layer
Weight Regularization	L2 (1×10^{-4})	Applied to dense layers
Activation Functions	Leaky ReLU, ReLU	Enables nonlinear feature learning
Output Activation	Linear	Continuous coordinate regression
Input Normalization	Batch normalization	Applied separately to RSSI and FTM
Train/Test Split	80% / 20%	Dataset partitioning
Framework	TensorFlow / Keras	Deep learning implementation

Note: RSSI = Received Signal Strength Indicator; FTM = Fine Timing Measurement; ReLU = Rectified Linear Unit.

4. RESULTS AND DISCUSSION

4.1 Training and validation performance

The model was trained for 750 epochs to get the best performance in terms of the evaluation metrics as shown in Figures 3-7. The first pivotal metric is the loss function (MSE) performance during the training phase. The results show a significant reduction in localization error between true and estimated positions (Figure 3). This decrease indicates that the model gradually learns the mapping between the fused signals and the actual position during training.

The second evaluation metric employed is the Cosine Similarity in the training phase, (Figure 4). The curves show a rapid increase in similarity during the early training epochs followed by a stabilized similarity ratio indicating that the model reaches its optimum learning capacity. Additionally, the curves show the capability of the model to overcome the overfitting that was observed throughout early training epochs. The initial rapid growth and subsequent leveling off suggest that the model learns fast to model the predominant fused signal-position dependency, and then makes finer adjustments in subsequent epochs.

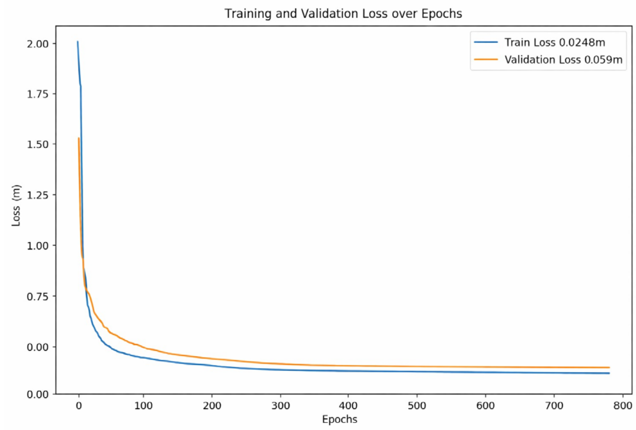


Figure 3. Training and validation loss over epochs

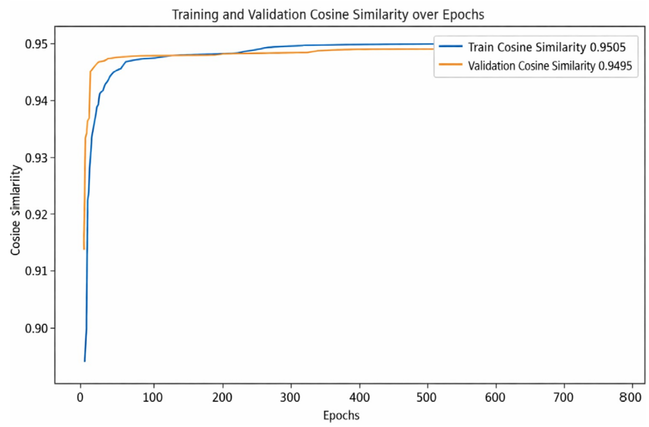


Figure 4. Training and validation cosine similarity over epochs

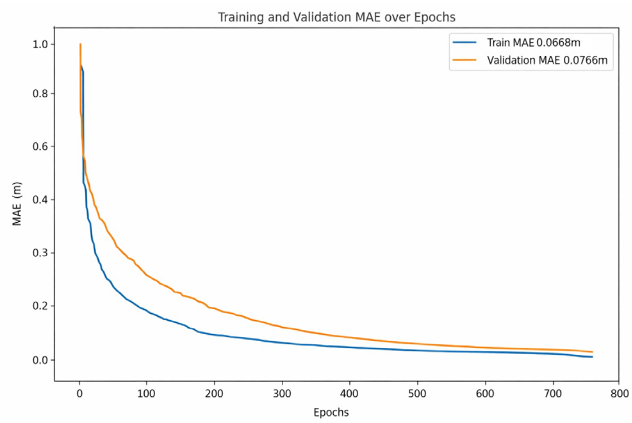


Figure 5. Training and validation mean absolute error (MAE) over epochs

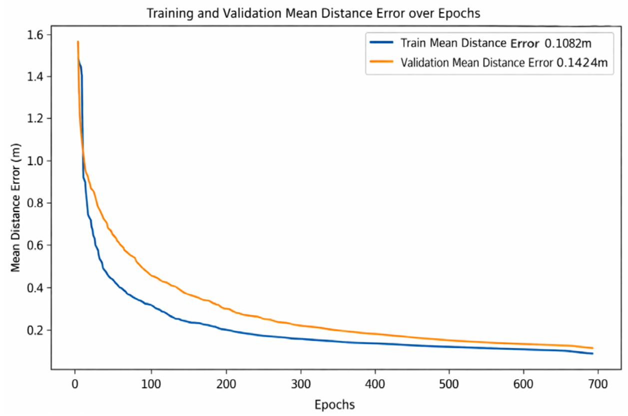


Figure 6. Training and validation mean distance error over epochs

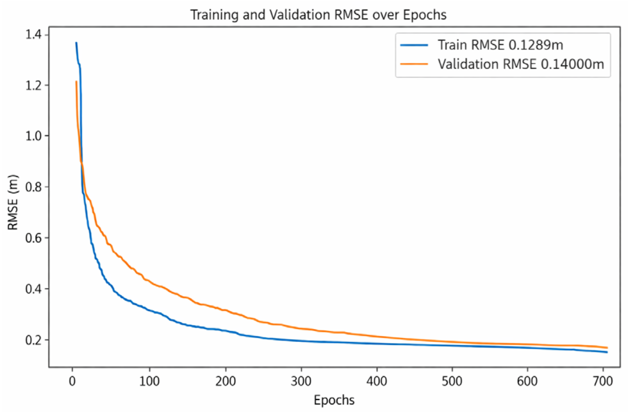


Figure 7. Training and validation root mean squared error (RMSE) over epochs

Another evaluation metric investigated is the mean absolute error (MAE) is shown in Figure 5. This metric can give us an intuitive and robust comprehensive measurement of the model accuracy without over-emphasizing large deviations by calculating the average magnitude of errors. The distinct trends of training and validation curves show superior levels of generalization. The difference between the training and validation curves is small, suggesting that the normalization and regularization procedures are effective in reducing overfitting, and that the accuracy is maintained on unseen data.

The Euclidean distance (Figure 6) measures the mean distance between true and estimated 2D locations, which exhibits a similar trend to the Cosine Similarity measure. Consequently, this evaluation metric reinforces the assumptions drawn from the previous metrics about the model's validity to address indoor localization tasks, as well as its ability to generalize.

The final metric used is the root mean squared error (RMSE) to assess the sensitivity and interpretability (Figure 7). Specifically, RMSE squares the errors between true and estimated values before averaging them, thereby assigning a higher weight given to large errors. This, in turn, forces the model to adjust its gradients to minimize large errors. The low and stable RMSE value means that there are few large positioning errors, which is crucial considering the signal fluctuation that occurs in an indoor environment. This metric can give interpretable measurements because it gives values in their original units (m) unlike MSE, which gives squared units (m^2).

The obtained results over the evaluation metrics show the ability of the model to obtain excellent estimations for indoor localization with strong generalization.

4.2 Test set evaluation

After completing the training of the proposed model on the training sub-dataset, a test for generalization was conducted on the unseen sub-dataset. The results were as follows: the loss function yielded a value of (0.0113 m^2), MAE (0.0297 m), RMSE (0.0456 m), Mean Distance Error (0.0477 m), and Cosine Similarity (0.977). The above results show the high ability of the model to converge and generalize when trained and tested on the ESP32C3-RSSI-FTM dataset. Furthermore, the results prove the capability of the implemented regularization methods and normalization techniques to ensure the stability of the learning process.

5. CONCLUSION AND FUTURE WORK

This research has introduced a dual-modality learning architecture designed to address indoor localization challenges. Specifically, it leverages a joint collaboration between RSSI and FTM signals acquired from ESP32C3 modules. By transforming nonhomogeneous RSSI and FTM signal ranges into a homogeneous stream, the model attains a unified feature map of signal strength and the time-of-flight information. The experiments conducted on the ESP32C3-RSSI-FTM dataset show excellent performance and strong ability to generalize in training and testing phases. In the testing phase, the model scores an MAE of 0.0297 m, an RMSE of 0.0456 m, and a mean distance error of 0.0477 m, with a high cosine similarity of 0.977. Finally, the present study has two main limitations. First, the model was trained and tested with one dataset, which was gathered in one building, and generalization of the model to other environments is still to be verified. Second, the proposed system is based on two-dimensional localization and does not yet support floor-level localization. These two points set the directions for our future work. An extensive comparison of the proposed model with external baseline approaches (benchmarking) on the same data set is envisaged for future work, and all the methods will be compared on a common coordinate system to enable a fair and uniform comparison. The framework will be expanded to support floor-level estimation using techniques like transfer learning to limit the need for extensive fingerprint acquisition. Additionally, more sensing modalities, like an inertial measurement unit (IMU), or geomagnetic signals will be embedded to enhance the robustness and accuracy of the model. This approach will be most beneficial in complex and congested environments.

REFERENCES

[1] Nguyen, S.M., Le, D.V., Havinga, P.J.M. (2024). Seeing the world from its words: All-embracing Transformers for fingerprint-based indoor localization. Pervasive and Mobile Computing, 100: 101912. https://doi.org/10.1016/j.pmcj.2024.101912

[2] Obeidat, H., Shuaieb, W., Obeidat, O., Abd-Alhameed, R. (2021). A review of indoor localization techniques and wireless technologies. Wireless Personal

Communications, 119: 289-327. https://doi.org/10.1007/s11277-021-08209-5

[3] Rizk, H., Elmogy, A., Yamaguchi, H. (2022). A robust and accurate indoor localization using learning-based fusion of Wi-Fi RTT and RSSI. Sensors, 22(7): 2700. https://doi.org/10.3390/s22072700

[4] Alaasam, H.A., Hamidkhani, M., Bachache, N.K., Soltan Aghaei, M. (2025). Improving the accuracy of indoor object localization in the Internet of Things using recurrent neural networks (LSTM) in challenging environments. IEEE Access, 13: 112446-112460. https://doi.org/10.1109/ACCESS.2025.3581417

[5] Singh, N., Choe, S., Punmiya, R. (2021). Machine learning based indoor localization using Wi-Fi RSSI fingerprints: An overview. IEEE Access, 9: 127150-127174. https://doi.org/10.1109/ACCESS.2021.3111083

[6] Isaia, C., Michaelides, M.P. (2023). A review of wireless positioning techniques and technologies: From smart sensors to 6G. Signals, 4(1): 90-136. https://doi.org/10.3390/signals4010006

[7] Kargar-Barzi, A., Farahmand, E., Taheri Chatrudi, N., Mahani, A., Shafique, M. (2024). An edge-based WiFi fingerprinting indoor localization using convolutional neural network and convolutional auto-encoder. IEEE Access, 12: 85050-85060. https://doi.org/10.1109/ACCESS.2024.3412676

[8] Hailu, T.G., Guo, X.S., Si, H.N. (2025). Indoor positioning systems as critical infrastructure: An assessment for enhanced location-based services. Sensors, 25(16): 4914. https://doi.org/10.3390/s25164914

[9] Kim, D., Park, J.H., Suh, Y.J. (2025). A Wi-Fi fingerprinting indoor localization framework using feature-level augmentation via variational graph auto-encoder. Electronics, 14(14): 2807. https://doi.org/10.3390/electronics14142807

[10] Alhmiedat, T. (2023). Fingerprint-based localization approach for WSN using machine learning models. Applied Sciences, 13(5): 3037. https://doi.org/10.3390/app13053037

[11] Chia, Z.Y., Goh, P.Y., Ong, L.Y., Tan, S.C. (2025). The challenge of dynamic environments in regard to RSSI-based indoor Wi-Fi positioning—A systematic review. Future Internet, 17(12): 540. https://doi.org/10.3390/fi17120540

[12] Yang, J.H., Wang, Y., Cheng, W., et al. (2024). A fingerprint dictionary processing approach in indoor localization system based on Wi-Fi. Scientific Reports, 14: 24944. https://doi.org/10.1038/s41598-024-75306-3

[13] Wu, Y., He, M.Y., Li, W., et al. (2025). Wi-Fi fine time measurement–Principles, applications, and future trends: A survey. Information Fusion, 118: 102992. https://doi.org/10.1016/j.inffus.2025.102992

[14] Xu, S.L., Wang, Y.J., Si, M.H. (2022). A two-step fusion method of Wi-Fi FTM for indoor positioning. Sensors, 22(9): 3593. https://doi.org/10.3390/s22093593

[15] Rathnayake, R.M.M.R., Maduranga, M.W.P., Tilwari, V., Dissanayake, M.B. (2023). RSSI and machine learning-based indoor localization systems for smart cities. Eng, 4(2): 1468-1494. https://doi.org/10.3390/eng4020085

[16] Sun, D.S., Wei, E.H., Ma, Z.X., Wu, C.X., Xu, S.Y. (2021). Optimized CNNs to indoor localization through BLE sensors using improved PSO. Sensors, 21(6): 1995. https://doi.org/10.3390/s21061995

[17] Njima, W., Bazzi, A., Chafii, M. (2022). DNN-based indoor localization under limited dataset using GANs and semi-supervised learning. IEEE Access, 10: 69896-69909. https://doi.org/10.1109/ACCESS.2022.3187837

[18] Yoon, J.H., Kim, H.J., Lee, D.S., Kwon, S.K. (2024). Indoor positioning method by CNN-LSTM of continuous received signal strength indicator. Electronics, 13(22): 4518. https://doi.org/10.3390/electronics13224518

[19] Guo, X.S., Ansari, N., Hu, F.Z., Shao, Y., Nkrow, R.E., Li, L. (2020). A survey on fusion-based indoor positioning. IEEE Communications Surveys & Tutorials, 22(1): 566-594. https://doi.org/10.1109/COMST.2019.2951036

[20] Yuen, B., Bie, Y.F., Cairns, D., Harper, G., Xu, J., Chang, C. (2022). Wi-Fi and Bluetooth contact tracing without user intervention. IEEE Access, 10: 91027-91044. https://doi.org/10.1109/ACCESS.2022.3201645