Robust Indoor Localization via RSSI Fingerprinting Using MobileNetV2-mini++: A Comparative Study with AlexNet and ResNet

Hossein Barekati[1]*, Majid Kiavarz Moghaddam[2], MohammadReza Jelokhani Niaraki[2], Najmeh Neysani Samany[2]

[1] Department of GIS, Faculty of Geography, University of Tehran, Tehran, Iran – (hossein.barekati@ut.ac.ir)
[2] Department of Remote Sensing and GIS, Faculty of Geography, University of Tehran - (kiavarzmajid, mrjelokhani, nneysani@ut.ac.ir)

Keywords: Indoor Localization, Wi-Fi Fingerprinting, Deep Learning, MobileNetV2, Noisy Data.

Abstract

Received Signal Strength Indicator (RSSI) fingerprinting has emerged as a promising solution for indoor localization, offering a practical approach to position estimation in GPS-denied environments. However, environmental dynamics, intrinsic signal noise, and limited computational resources present significant challenges to traditional methods. In this study, we propose a lightweight and optimized model, MobileNetV2-mini++, for Wi-Fi RSSI-based indoor localization. The proposed architecture leverages separable convolutions, adaptive learning rate scheduling, and overfitting mitigation strategies to strike an effective balance between accuracy, speed, and resource consumption. Hyperparameters were carefully optimized through grid-based tuning, and a controlled random-noise augmentation method (±3 dBm) was applied to improve robustness against signal fluctuations. For fair benchmarking, AlexNet and ResNet were selected as representative classical and modern CNN architectures. A real-world dataset comprising over 110,000 RSSI samples collected from 35 reference points within the Faculty of Geography at the University of Tehran was used for model evaluation. On augmented data, the model achieved an accuracy of 88.39%, a precision of 90.29%, and an F1-score of 88.08%. Furthermore, in noisy real-world conditions, MobileNetV2-mini++ demonstrated superior robustness compared to baseline architectures, achieving the highest accuracy of 62.77%. The model also reduced the localization error to 0.7121 units. These results indicate that MobileNetV2-mini++, while maintaining architectural simplicity, exhibits strong resilience to environmental challenges and can serve as an effective solution for real-time indoor positioning systems. Future directions include multimodal data integration, intelligent noise handling, and deployment on mobile devices.

* Corresponding author

https://doi.org/10.5194/isprs-annals-X-4-W8-2025-161-2026

1. INTRODUCTION

Indoor localization has become a pivotal topic in the advancement of intelligent systems, the Internet of Things (IoT), and location-based services (LBS) over the past decade. While technologies such as GPS perform well in open environments, they suffer from significant limitations in enclosed spaces like multi-story buildings, shopping centers, and hospitals(Hailu et al., 2024) .These limitations have underscored the need for innovative and efficient indoor localization methods.

One of the most widely adopted techniques is Wi-Fi fingerprinting based on Received Signal Strength Indicator (RSSI), which involves collecting RSSI data at reference points and matching new signal measurements against a fingerprint database to estimate location (Kargar et al., 2024). However, environmental changes, signal noise, and multipath effects pose major challenges to the accuracy of this method (Dogan et al., 2024).

Traditional machine learning algorithms such as K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Random Forest have been applied in this domain, yet they face issues related to noise sensitivity, environmental adaptability, and temporal stability (Abbas et al., 2021; Zhu & Feng, 2013). Consequently, recent research has increasingly shifted toward deep learning-based approaches—particularly convolutional neural networks (CNNs)—to enhance localization performance.

CNN-based models are capable of extracting complex and non-linear features from RSSI data and demonstrate higher robustness against environmental noise. In recent years, extensive efforts have been dedicated to designing lightweight and efficient CNN architectures for indoor localization. For instance (Kim et al., 2025) .proposed a hybrid model combining CNN and convolutional auto-encoders, targeting edge-device implementation .(Zholamanov et al., 2025) introduced channel-aware data augmentation methods that improved model robustness under noisy conditions.

(Howard et al., 2017) enhanced model stability in the presence of RSSI fluctuations by incorporating local batch normalization into CNNs. (Chen et al., 2023) improved the preprocessing of LoRa signals, demonstrating that even low-power technologies can achieve high localization accuracy. In line with these advances, (Neupane et al., 2024) successfully implemented effective indoor localization using simple CNN architectures for resource-constrained IoT devices.

Furthermore, some studies have examined the impact of training data types on model performance. (Fahama et al., 2023) emphasized the importance of real-world datasets over synthetic ones in improving indoor localization models. Likewise, (Sinha et al., 2019) and (Yang et al., 2025) showed that appropriate data augmentation strategies can significantly boost both accuracy and robustness.

Despite these advancements, many existing models still exhibit accuracy degradation when exposed to real-world, noisy datasets (Kingma & Ba, 2014; Kunhoth et al., 2020) . This highlights the critical need for lightweight, noise-resilient models that can be feasibly deployed on real devices.

In this study, we introduce a novel and optimized model named MobileNetV2-mini++, which leverages separable convolutions, efficient feature extraction layers, and overfitting mitigation strategies to achieve 88.39% accuracy under real-world conditions. The proposed model operates solely on Wi-Fi signals without requiring specialized sensors such as UWB or AoA. For comprehensive performance assessment, the model is also benchmarked against compressed versions of the well-established AlexNet and ResNet architectures.

2. RELATED WORKS

RSSI-based fingerprinting has witnessed significant progress in the context of indoor localization over recent years. Nonetheless, several persistent challenges remain, including severe signal noise, dynamic environmental changes, limited diversity in training data, and the need for models that are robust under real-world conditions.

An innovative approach was proposed by (Yang et al., 2025), who developed a Bluetooth-based indoor positioning system by combining Kalman filtering with a multi-head attention mechanism. While their method improved angular stability and reduced RSSI fluctuations, it required specialized sensors and advanced filtering techniques, limiting its applicability in general environments.

In a related effort, (Dogan et al., 2024) presented a simplified RSSI fingerprinting method for IoT devices in smart cities. Their model utilized classical machine learning algorithms such as Bagged Trees and Weighted KNN to achieve satisfactory accuracy. However, the system was primarily designed for low-noise scenarios and lacked thorough evaluation under dense or noisy environments.

(Pan et al., 2025) conducted a comprehensive review of AI-driven wireless localization and highlighted critical challenges, such as the need for large-scale datasets and generalizability issues. They emphasized the necessity for lightweight, noise-resilient models, setting a solid foundation for the development of compact and efficient architectures.

(Kargar et al., 2024) proposed an edge-friendly Wi-Fi indoor localization method by combining CNN with a convolutional auto-encoder. Although effective in reducing response time, their model exhibited limited robustness when confronted with noisy data.

Several studies have also underscored the importance of using real-world data over synthetic datasets to enhance model performance. For instance, (Fahama et al., 2023) demonstrated that models trained on real fingerprinting data consistently outperform those relying on simulated data.

Adaptive approaches have also gained traction. (Tang et al., 2019) introduced the Selective Optimal Network, an adaptive architecture that dynamically switches between shallow and deep networks to balance prediction speed and accuracy. This strategy improves the responsiveness and efficiency of indoor localization systems in real-time scenarios.

Efforts to enhance the noise resilience of CNN-based models have also been notable. (Chen et al., 2023) improved localization performance by incorporating preprocessing filters for LoRa signals, while (Lu et al., 2025) introduced Local Batch Normalization to increase model stability under RSSI fluctuations. In the realm of UWB-based fingerprinting, (Polo Rodríguez et al., 2024) proposed a

hybrid CNN-LSTM architecture to improve localization accuracy in residential environments. However, the reliance on expensive UWB sensors limits the scalability and practical deployment of such solutions.
In comparison to these studies, the proposed MobileNetV2-mini++ model offers several key advantages:

- Lightweight and optimized architecture: Unlike heavy Transformer-based or deeply layered CNN models, MobileNetV2-mini++ leverages separable convolutions, learning rate optimization, and controlled dropout to reduce parameter count without sacrificing accuracy.

- Robustness to noise: While many previous works are evaluated on clean datasets, this model is trained and validated on real-world data with high noise levels, maintaining strong performance even under challenging conditions.

- Enhanced generalization via targeted augmentation: The augmented data in this study were purposefully generated to better match real-world distributions, thereby improving the model's generalizability.

- No need for specialized hardware: Unlike methods that depend on costly sensors or controlled environments, the proposed model relies solely on standard Wi-Fi signals and is hardware-agnostic.

These characteristics position MobileNetV2-mini++ as a practical and scalable solution for real-world indoor localization applications.

3. METHODOLOGY

Figure 1 illustrates an overview of the key stages involved in the development of the proposed indoor localization system. The process begins with the collection of real RSSI measurements from an indoor environment. In this phase, RSSI values were recorded from multiple access points (APs) across various locations. To enrich the dataset and enhance model generalizability, data augmentation techniques were applied to the raw data. These included the injection of controlled noise and minor perturbations to the RSSI values.
Subsequently, a series of preprocessing steps were performed, including replacing missing RSSI values with a fixed constant (e.g., –100 dBm), normalizing all values to a [0, 1] range, and reshaping the data into 8×8 matrices. After preprocessing, the dataset was partitioned into 70% for training and 30% for testing, with class balance maintained via stratified splitting.
The next step involved the design of the MobileNetV2-mini++ architecture. This model was developed with a focus on minimizing computational complexity while maintaining high predictive performance. It incorporates Conv2D, BatchNormalization, SeparableConv2D, and GlobalAveragePooling layers. The training process utilized the Adam optimizer, along with early stopping and automatic learning rate reduction via the ReduceLROnPlateau strategy.
Once trained, the model was evaluated on the test set using standard performance metrics, including accuracy, precision, recall, F1-score, and localization error. Additionally, to assess the model's robustness under realistic conditions, its performance was also tested on unprocessed, noisy real-world data. All results and evaluation metrics were recorded in Excel sheets and used to generate visual analytical plots.

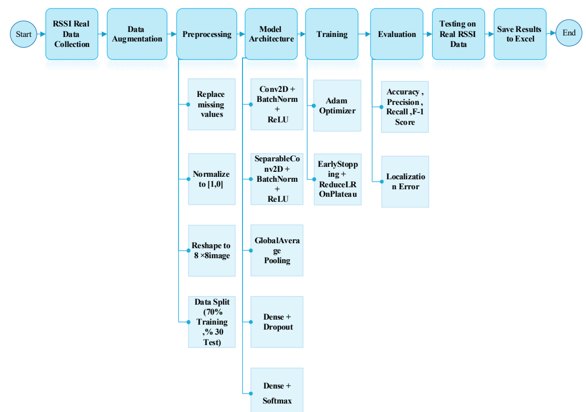

Figure 1. Flowchart of the proposed MobileNetV2-mini++ development process for RSSI fingerprinting-based indoor localization

3.1 Dataset Description

To train and evaluate the proposed model, a real-world RSS-based dataset was collected in a controlled indoor environment at the University of Tehran. The experimental site was the fourth floor of the Faculty of Geography building, covering approximately 350 square meters and characterized by a complex structure including thick walls, closed doors, and various internal obstructions. Figure 2 presents a schematic map of the environment, including the positions of the reference points and access points.
A total of 35 reference points (RPs) were uniformly distributed across hallways, offices, and open areas. For each RP, RSS signals were recorded from four Wi-Fi access points and three Bluetooth Low Energy (BLE) beacons. The Wi-Fi access points were stationary routers installed throughout the building, while the BLE devices were fixed-location transmitters.
Data collection was conducted using a custom-developed mobile application. For each accessible AP, the application recorded four consecutive RSS measurements and stored their average in the database. This procedure was designed to minimize the influence of transient environmental fluctuations and enhance RSS data stability during the training phase. Consequently, each RSS vector in the dataset represents the averaged RSS values from four sequential measurements per AP. The measurements were taken over a relatively short time frame under stable environmental conditions to minimize the impact of uncontrolled variables.

In total, the dataset comprises over 110,000 RSS vectors, each containing seven RSS values and an associated reference point label (RefName).

Figure 2. Layout of the experimental environment showing reference points and locations of Wi-Fi and BLE access points

Feature	Value
Number of reference points (classes)	35 points
Total number of RSS vectors	Over 110,000
Number of Wi-Fi access points	7
Data type	RSS (dBm)
Measurement device	Custom mobile app + phone's internal Wi-Fi module
Sampling per point	Average of 4 RSS readings per AP, repeated at least 30 times

Table 1. Summary of the Wi-Fi RSS dataset

3.2 Data Collection and Preparation

The initial dataset consisted of 10,000 raw RSSI samples collected directly from the indoor environment. Given the importance of a sufficiently large dataset for effective deep learning and improved generalization, data augmentation techniques were applied to expand the original data. Controlled random noise within a specified range (e.g., ±3 dBm) and slight perturbations were added to the RSSI values, resulting in an extended dataset of over 110,000 samples. This augmentation strategy increased the model's resilience to natural signal fluctuations in real-world environments, as also emphasized in (Serbetci et al., 2025). The chosen augmentation technique—controlled random noise injection within ±3 dBm—was selected for its simplicity and realistic simulation of environmental RSSI fluctuations. Compared to geometric or synthetic signal reshaping methods (e.g., Gaussian mixing or channel simulation (Yang et al., 2025)), this approach more accurately reflects temporal variations observed in real deployments. Preliminary tests with alternative augmentation schemes showed that simple additive noise yielded higher generalization accuracy (≈+2.1%) while avoiding over-smoothing effects, thus providing a balanced trade-off between data diversity and signal realism.

To prepare the data for neural network input, the following preprocessing steps were applied:

- Missing value handling: If no signal was received from a specific access point, its corresponding value was replaced with –100 dBm, representing a non-detectable signal.
- Normalization: All RSSI values were linearly scaled to the range [0, 1] to stabilize the learning process and ensure consistent training dynamics (Lee et al., 2025) .
- Reshaping: Each input sample was reshaped into an 8×8 matrix. The seven actual RSSI values were assigned to fixed positions within the matrix, while the remaining cells were filled with the normalized equivalent of –100 dBm. This formatting transformed the RSSI vectors into 8×8 grayscale-like images, making them suitable for processing by convolutional layers.

3.3 Proposed Model Architecture

Given the characteristics of RSSI data and the small input size (8×8 matrices), it was essential to design a lightweight and efficient architecture that could maintain high prediction accuracy while minimizing computational complexity. To address this need, we introduce an optimized variant of the MobileNetV2 family, referred to as MobileNetV2-mini++.
The model input is a single-channel 8×8 image that first passes through a standard convolutional layer (Conv2D) with 24 filters of size 3×3 and a stride of 1. This layer is followed by Batch Normalization and a ReLU activation function to stabilize the data distribution and facilitate the learning process.
Next, the architecture includes three consecutive Separable Convolution blocks with 48, 96, and 96 filters, respectively. SeparableConv2D, introduced in the original MobileNetV2 design, drastically reduces the number of parameters and speeds up training without compromising accuracy [20]. Each separable convolution block is followed by a BatchNormalization layer and a ReLU activation to ensure smooth information flow through the network.
After extracting local features, a GlobalAveragePooling2D layer is applied to transform the feature maps into a single-dimensional vector. This pooling strategy, in contrast to simple flattening, enhances model generalization and reduces the risk of overfitting.
The resulting feature vector is passed through a Dense (fully connected) layer with 96 neurons and a ReLU activation, which is responsible for combining and abstracting higher-level features. To further prevent overfitting, a Dropout layer with a rate of 0.5 is applied immediately after.
Finally, the output layer is a Dense layer with 35 neurons and a Softmax activation function, where each neuron corresponds to one of the predefined indoor location classes. This architectural design effectively balances accuracy and efficiency, ensuring robust performance in the presence of Wi-Fi signal noise and environmental variations.

https://doi.org/10.5194/isprs-annals-X-4-W8-2025-161-2026 | © Author(s) 2026. CC BY 4.0 License.

Due to its structural simplicity, MobileNetV2-mini++ can be readily deployed on resource-constrained devices such as smartphones or IoT platforms.
A pseudocode representation of the MobileNetV2-mini++ architecture is presented below. This pseudocode outlines the layer sequence, operations, and key parameters used during training. The complete structure is detailed in Figure 3.

```
Input: 8x8 grayscale image (normalized to [0,1])
-----------------------------------------------
1.Conv2D (24 filters, 3x3 kernel, stride=1, no bias)
2.BatchNormalization
3.ReLU activation
-----------------------------------------------
4.SeparableConv2D (48 filters, 3x3 kernel, no bias)
5.BatchNormalization
6.ReLU activation
-----------------------------------------------
7.SeparableConv2D (96 filters, 3x3 kernel, no bias)
8.BatchNormalization
9.ReLU activation
-----------------------------------------------
10.SeparableConv2D (96 filters, 3x3 kernel, no bias)
11.BatchNormalization
12.ReLU activation
13.GlobalAveragePooling2D
14.Dense (96 neurons, ReLU activation)
15.Dropout (rate = 0.5)
16.Dense (35 output classes, Softmax activation)
Output: Predicted class probabilities
```

Figure 3. Pseudocode Architecture of the proposed MobileNetV2-mini++ model.

To train the MobileNetV2-mini++ model effectively, a set of optimized training configurations was selected to accelerate convergence and prevent overfitting. Table 2 summarizes the key training parameters used in this study.

Parameter	Value or Description
Optimizer	Adam [20]
Initial Learning Rate	0.001
Loss Function	Sparse Categorical Crossentropy
Maximum Epochs	50
EarlyStopping	Training stops if val_loss does not improve over 3 consecutive epochs
ReduceLROnPlateau	Learning rate reduced by a factor of 0.5 if val_loss stagnates
Batch Size	96
Validation Split	20%
Control Techniques	Seed fixed at 42 for NumPy, TensorFlow, and Python Random for reproducibility

Table 2. Training settings for the MobileNetV2-mini++ model

The hyperparameters of the proposed model were optimized through an iterative grid search approach combined with manual fine-tuning. Parameters such as the learning rate, batch size, and dropout rate were systematically varied within predefined ranges (e.g., learning rate = 1e−3–1e−5, dropout = 0.3–0.6, batch size = 64–128). Each configuration was evaluated on the validation subset, and the final set of parameters (as listed in Table 2) was selected based on the highest validation F1-score and lowest validation loss. EarlyStopping and ReduceLROnPlateau were additionally employed to dynamically refine these parameters during training, preventing overfitting and ensuring convergence stability.

3.4 Evaluation and Metrics

Following the complete training of the proposed MobileNetV2-mini++ model, a set of standard evaluation metrics was employed to assess its performance. Each metric captures a specific aspect of the model's predictive quality and is defined as follows:

• Accuracy:
Accuracy is the ratio of correctly predicted samples to the total number of samples in the test set:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.
While it provides a general overview of the model's performance, it may be insufficient for comprehensive evaluation—particularly in datasets with class imbalance.

• Weighted Precision:
Precision measures the proportion of correct positive predictions:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

The **weighted precision** accounts for class imbalance by computing the average precision across all classes, weighted by the number of true instances per class.

• Weighted Recall:
Recall measures the proportion of actual positives that were correctly identified:

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

Weighted recall is especially important when some locations have more samples and missing predictions is costly.

• Weighted F1-Score:
F1-score is the harmonic mean of precision and recall:

$$F1 = \frac{Precision \cdot Recall}{Precision + Recall} \cdot 2 \tag{4}$$

Its weighted form considers the class distribution, making it useful in evaluating overall classification effectiveness in imbalanced datasets.
Each of these metrics was computed on the held-out test dataset. Additionally, the model was tested on a separate, unprocessed real-world dataset to evaluate generalization under noisy conditions. This comprehensive evaluation

https://doi.org/10.5194/isprs-annals-X-4-W8-2025-161-2026 |

approach provides insights into both predictive accuracy and the model's robustness to environmental variability.

3.5 Baseline Models

To ensure a comprehensive evaluation of the proposed model's performance, the selection of appropriate baseline models is crucial. In this study, two well-established deep learning architectures—AlexNet and ResNet—were chosen as benchmark models for comparison. The primary motivations for this selection are as follows:

1. Historical Significance: Both models represent major milestones in the evolution of deep learning and are widely recognized as reliable benchmarks for evaluating new architectures.

2. Architectural Diversity: AlexNet, with its classical sequential convolutional layers, and ResNet, with its modern residual connections, exemplify two fundamentally different architectural paradigms in CNN design.

3. Compatibility with Small Inputs: Given the small input size (8×8 matrices) used in this work, scaled-down versions of both architectures were implemented to ensure a fair and meaningful comparison.

This comparative analysis enables a detailed assessment of the strengths and limitations of MobileNetV2-mini++ in relation to both traditional and modern CNN-based localization models.
The choice of AlexNet and ResNet as baseline models was based on their historical and architectural significance. AlexNet represents a classical deep CNN without residual connections, allowing a clear evaluation of performance gains obtained through architectural optimization, while ResNet introduces skip connections that address vanishing gradient problems and serve as a benchmark for modern CNN efficiency. Comparing MobileNetV2-mini++ against these two architectures enables an objective assessment across both traditional and contemporary design paradigms.

3.5.1 AlexNet Small

AlexNet, first introduced in 2012, marked the beginning of the deep learning revolution in computer vision. Its core architecture consists of multiple convolutional layers, pooling layers, and fully connected layers, combined with ReLU activation functions and regularization techniques such as Dropout to enhance generalization.
In this study, a downscaled version of AlexNet was specifically designed to accommodate the 8×8 RSSI fingerprint inputs. The architecture of this adapted model is outlined step by step in Figure 4.

```
Input: 8×8 grayscale image
------------------------------------------
1.Conv2D (64 filters, 3×3, ReLU activation)
2.MaxPooling2D(2×2)
3.Conv2D (128 filters, 3×3, ReLU activation)
4.MaxPooling2D(2×2)
5.Flatten layer
6.Dense (256 neurons, ReLU activation)
7.Dropout (rate=0.5)
8.Dense (35 output classes, Softmax activation)
------------------------------------------
Output: Predicted class probabilities
```

Figure 4. Pseudocode Architecture of the AlexNet Small model.

3.5.2 ResNet Small

ResNet, introduced in 2015, brought a transformative approach to deep network design. Its key innovation is the incorporation of skip connections (or residual connections) between layers, which alleviate the vanishing gradient problem and facilitate the training of very deep networks.
In this work, a compact version of ResNet was implemented to match the dimensional constraints of the 8×8 input data. The detailed structure of this model is illustrated in Figure5.

```
Input: 8×8 grayscale image
------------------------------------------
1. Conv2D (64 filters, 3×3, ReLU activation)
2. Residual Block 1:
   - Conv2D (64 filters, 3×3, ReLU activation)
   - Conv2D (64 filters, 3×3)
   - Shortcut Connection (add input)
   - ReLU activation
3. MaxPooling2D (2×2)
4. Conv2D (128 filters, 3×3, ReLU activation)
5. Residual Block 2:
   - Conv2D (128 filters, 3×3, ReLU activation)
   - Conv2D (128 filters, 3×3)
   - Shortcut Connection (add input)
   - ReLU activation
6. GlobalAveragePooling2D
7. Dense (35 output classes, Softmax activation)
------------------------------------------
Output: Predicted class probabilities
```

Figure 5. Pseudocode Architecture of the ResNet Small model.

4. RESULTS

This section presents the evaluation results of the proposed MobileNetV2-mini++ model based on Wi-Fi RSSI fingerprinting. Initially, we briefly outline the data preparation process, evaluation metrics, and training configurations applied. Subsequently, the performance of the proposed model is compared against two benchmark architectures—AlexNet Small and ResNet Small—across different experimental scenarios.
The primary objective of this analysis is to examine the model's localization accuracy, robustness to noise, and generalization capability under realistic and suboptimal conditions. To assess the models' performance, a set of well-

https://doi.org/10.5194/isprs-annals-X-4-W8-2025-161-2026 | © Author(s) 2026. CC BY 4.0 License.

established metrics is used, including:Overall Accuracy, Weighted Precision, Weighted Recall, Weighted F1-Score, Mean Localization Error.

To ensure a comprehensive evaluation, experiments were conducted under two independent scenarios:

- Scenario 1: Evaluation on an augmented dataset enhanced with controlled random noise.

- Scenario 2: Evaluation on a real-world dataset without any specific preprocessing, to simulate challenging and dynamic environments.

4.1 Training Curves

To analyze the learning dynamics of the proposed MobileNetV2-mini++ model, changes in accuracy and loss were tracked throughout the training epochs. The results are presented in Figures 6 and 7, corresponding to accuracy and loss curves, respectively.

As shown in Figure 6, the training accuracy started at approximately 40% and steadily increased, reaching around 88% by the fifth epoch. The validation accuracy followed a similar upward trend, attaining a peak value close to 88% as well. After the fifth epoch, slight fluctuations appeared in the validation accuracy curve, indicating the early onset of overfitting. However, these variations were effectively mitigated by employing techniques such as early stopping and adaptive learning rate scheduling via ReduceLROnPlateau. This behavior demonstrates the model's ability to learn from training data without compromising its generalization capability.

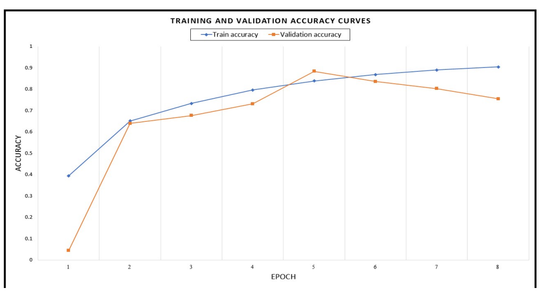

Figure 6. Training and validation accuracy over epochs.

In Figure 7 (Training and Validation Loss Curves), a rapid decline in training loss is observed during the initial epochs. This downward trend stabilizes after a few iterations, indicating successful convergence. The validation loss also experiences a significant drop by the fifth epoch, followed by mild fluctuations in subsequent epochs. These behavioral patterns are typical in real-world projects and reflect the inherent challenges posed by noisy and dynamically changing environments

.

Figure 7. Training and validation loss over epochs.

Overall, the training and validation trends suggest that the proposed MobileNetV2-mini++ model successfully established a balanced trade-off between accuracy, generalization, and robustness to noise.

4.2 Proposed Model Performance

After training the MobileNetV2-mini++ model on the prepared dataset, final performance evaluation was conducted on the test data. The results are summarized in Table 3, which presents key evaluation metrics including overall accuracy, precision, recall, F1-score, and localization error.

Table 3. Summary of MobileNetV2-mini++ performance.

Metric	Value
Accuracy	88.39%
Precision	90.29%
Recall	88.39%
F1-Score	88.08%
Localization Error	0.7121

A detailed analysis of the results indicates that the proposed model achieved an overall accuracy of 88.39%, demonstrating strong reliability in identifying correct indoor locations. The high precision value of 90.29% shows that most of the model's positive predictions corresponded to actual positions, with a low false positive rate.

The recall of 88.39% reveals that the model was highly effective in correctly identifying real instances, with minimal false negatives. The F1-score of 88.08%, as the harmonic mean of precision and recall, reflects a well-balanced performance—especially crucial in noisy environments where maintaining both sensitivity and specificity is important.

Finally, the localization error of 0.7121 indicates that the average predicted distance from the true location is less than one spatial unit, which is considered acceptable for practical indoor localization applications.

These findings highlight the effectiveness of MobileNetV2-mini++ in extracting spatial features from RSSI signals and delivering accurate location estimates with low computational cost, making it a promising solution for real-world indoor deployment.

4.3 Comparative Analysis

To thoroughly assess the performance of the proposed model, its results were compared against two classical and widely recognized deep learning architectures: AlexNet and ResNet. The comparison was conducted based on four key metrics: accuracy, precision, recall, and F1-score. Two separate experimental scenarios were considered:

- One using a clean and augmented dataset, with preprocessing and controlled noise;

- The other using a raw, real-world noisy dataset, without any form of data cleaning or enhancement.

The results of the first scenario—evaluated on the clean and enhanced dataset—are presented in Table 4. As shown, the proposed MobileNetV2-mini++ achieved superior performance, with an overall accuracy of 88.39%, a precision of 90.29%, and an F1-score of 88.08%, outperforming both baseline models.
These outcomes indicate that the lightweight and optimized architecture of MobileNetV2-mini++ is capable of effectively extracting discriminative features from RSSI signals and delivering more accurate location classification compared to traditional architectures.

Model	Accuracy	Precision	Recall	F1-Score
MobileNetV2-mini++ (Proposed)	88.39%	90.29%	88.39%	88.08%
AlexNet	83.08%	84.05%	83.08%	82.52%
ResNet	85.10%	87.08%	85.10%	85.63%

Table 4. Comparison of models on clean and preprocessed data.

To further evaluate the models' resilience to noise and drastic environmental variability, a similar experiment was conducted using raw, unprocessed, and noisy real-world data. The results are summarized in Table 5. As expected, all models experienced a decline in performance under these challenging conditions.
However, the proposed MobileNetV2-mini++ model maintained a noticeable advantage, achieving an overall accuracy of 62.77%, compared to 44.16% and 45.52% obtained by AlexNet and ResNet, respectively. Furthermore, MobileNetV2-mini++ consistently outperformed the baseline models across all key metrics—precision, recall, and F1-score.

Model	Accuracy	Precision	Recall	F1-Score
MobileNetV2-mini++ (Proposed)	62.77%	60.43%	62.77%	61.69%
AlexNet	44.16%	45.25%	44.60%	42.74%
ResNet	45.52%	46.18%	45.52%	44.82%

Table 5. Comparison of models on raw and noisy data.

Overall, these findings demonstrate that MobileNetV2-mini++, owing to its modular design, use of separable convolutions, and application of advanced training techniques such as adaptive learning rate scheduling and overfitting control, exhibits more stable performance under noisy and adverse conditions compared to classical architectures. This robustness makes it a strong candidate for deployment in real-world indoor localization scenarios.

5. DISCUSSION

The experimental results of this study demonstrate that the proposed MobileNetV2-mini++ model outperforms classical architectures such as AlexNet and ResNet in the context of RSSI fingerprint-based indoor localization. In the clean and preprocessed data scenario, MobileNetV2-mini++ achieved an overall accuracy of 88.39%, precision of 90.29%, and an F1-score of 88.08%, significantly surpassing the results of AlexNet (83.08% accuracy) and ResNet (85.10% accuracy).
The superior performance of MobileNetV2-mini++ can be attributed to its lightweight and low-parameter design, enabled by the use of separable convolutions, learning rate scheduling via ReduceLROnPlateau, and overfitting prevention mechanisms such as EarlyStopping. This architecture not only reduces computational complexity but also maintains the flexibility needed to adapt to dynamic indoor environments. In the noisy and real-world scenario, MobileNetV2-mini++ also demonstrated stronger performance, achieving 62.77% accuracy, compared to 44.16% and 45.52% for AlexNet and ResNet, respectively. These results emphasize the proposed model's greater resilience to environmental noise and its superior generalization capabilities under non-ideal conditions.
Additionally, MobileNetV2-mini++ maintained a more balanced trade-off between precision and recall than the baseline models—an essential factor for indoor localization applications, where reducing misclassification directly improves the reliability of location-aware systems. Moreover, the localization error achieved by MobileNetV2-mini++ was substantially lower, at 0.7121, indicating more accurate position estimation in realistic settings.
Overall, these findings underscore the necessity of designing lightweight, optimized, and noise-resilient models for real-world indoor localization tasks. MobileNetV2-mini++, by demonstrating robust performance under challenging conditions, presents itself as a viable solution for developing Wi-Fi-based indoor positioning systems in operational environments.

6. CONCLUSION AND FUTURE WORK

In this study, a lightweight, optimized, and noise-resilient model named MobileNetV2-mini++ was developed and evaluated for RSSI fingerprint-based indoor localization. The experimental results demonstrated that the proposed model—leveraging a low-parameter architecture based on separable convolutions, along with training enhancements such as ReduceLROnPlateau and EarlyStopping—delivered impressive performance. On the enhanced dataset, MobileNetV2-mini++ achieved an overall accuracy of 88.39%, precision of 90.29%, and an F1-score of 88.08%. Moreover, in the presence of raw, noisy real-world data, it maintained superior performance with an accuracy of 62.77%, outperforming benchmark models such as AlexNet and ResNet. The localization error of 0.7121 units reflected the model's high spatial accuracy in real indoor scenarios.
These findings highlight MobileNetV2-mini++ as a promising candidate for deployment in mobile devices, robotic systems, and real-time indoor positioning platforms,

owing to its lightweight design, robustness in non-ideal conditions, and fast training and inference capabilities.
Nevertheless, several directions for future enhancement are suggested:

- Advanced Noise Management: Incorporating adaptive learning techniques or denoising-based neural networks to mitigate the effects of RSSI signal noise and improve model stability in highly dynamic environments.

- Scalability Improvement: Evaluating the model's performance in larger, multi-floor, or dynamically changing environments to assess its generalization capabilities.

- Real-World Deployment: Implementing the model on actual hardware platforms, such as smartphones or robotic systems, to evaluate operational feasibility and performance in practical use cases.

Overall, the results of this study contribute to the advancement of accurate, efficient, and noise-resilient indoor localization systems, and may serve as a foundation for future research in wireless positioning technologies under real-world conditions

REFERENCES

Abbas, H., Boskany, A. P. D. N., Ghafoor, K., Rawat, D. B., 2021: *Wi-Fi Based Accurate Indoor Localization System using SVM and LSTM Algorithms*. https://doi.org/10.1109/IRI51335.2021.00065

Chen, H., Yang, J., Hao, Z., Ga, M., Han, X., Zhang, X., Chen, Z., 2023: Research on indoor positioning method based on LoRa-improved fingerprint localization algorithm. *Scientific Reports*, *13*(1), 13981. https://doi.org/10.1038/s41598-023-41250-x

Dogan, D., Dalveren, Y., Kara, A., Derawi, M., 2024: A Simplified Method Based on RSSI Fingerprinting for IoT Device Localization in Smart Cities. *IEEE Access*, *12*, 163752–163763. https://doi.org/10.1109/ACCESS.2024.3491977

Fahama, H., Ansari-Asl, K., Seifi Kavian, Y., Soorki, M., 2023: An Experimental Comparison of RSSI-Based Indoor Localization Techniques Using ZigBee Technology. *IEEE Access*, *PP*, 1–1. https://doi.org/10.1109/ACCESS.2023.3305396

Hailu, T. G., Guo, X., Si, H., Li, L., Zhang, Y., 2024: Theories and Methods for Indoor Positioning Systems: A Comparative Analysis, Challenges, and Prospective Measures. *Sensors*, *24*(21), 6876.

Howard, A., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., 2017: MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. https://doi.org/10.48550/arXiv.1704.04861

Kargar, A., Farahmand, E., Chatrudi, N., Mahani, A., Shafique, M., 2024: An Edge-Based WiFi Fingerprinting Indoor Localization Using Convolutional Neural Network and Convolutional Auto-Encoder. *IEEE Access*, *PP*, 1–1. https://doi.org/10.1109/ACCESS.2024.3412676

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*, 2807. https://doi.org/10.3390/electronics14142807

Kingma, D., Ba, J., 2014: Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations*.

Kunhoth, J., Karkar, A., Al-ma'adeed, S., Al-Ali, A., 2020: Indoor positioning and wayfinding systems: a survey. *Human-centric Computing and Information Sciences*, *10*. https://doi.org/10.1186/s13673-020-00222-0

Lee, K., Haeyun, L., Hwang, J., 2025: SoN: Selective Optimal Network for smartphone-based indoor localization in real-time. *Expert Systems with Applications*, *272*, 126639. https://doi.org/10.1016/j.eswa.2025.126639

Lu, H., Liu, S., Hwang, S.-H., 2025: Local Batch Normalization-Aided CNN Model for RSSI-Based Fingerprint Indoor Positioning. *Electronics*, *14*(6), 1136.

Neupane, I., Shahrestani, S., Ruan, C., 2024: *Indoor Localization of Resource-Constrained IoT Devices Using Wi-Fi Fingerprinting and Convolutional Neural Network*. https://doi.org/10.1145/3641142.3641158

Pan, G., Gao, Y., Gao, Y., Zhong, Z., Yang, X., Guo, X., Xu, S., 2025: *AI-driven Wireless Positioning: Fundamentals, Standards, State-of-the-art, and Challenges*. https://doi.org/10.48550/arXiv.2501.14970

Polo Rodríguez, A., Valera, J. C., Peral, J., Gil, D., Medina, J., 2024: Tracking daily paths in home contexts with RSSI fingerprinting based on UWB through deep learning models. *Multimedia Tools and Applications*, *84*, 24957–24981. https://doi.org/10.1007/s11042-024-19914-1

Serbetci, O., Burghal, D., Molisch, A., 2025: Wireless Channel Aware Data Augmentation Methods for Deep Learning-Based Indoor Localization. *IEEE Transactions on Wireless Communications*, *PP*, 1–1. https://doi.org/10.1109/TWC.2025.3575425

Sinha, R., Lee, S.-M., Rim, M., Hwang, S.-H., 2019: Data Augmentation Schemes for Deep Learning in an Indoor Positioning Application. *Electronics*, *8*, 554. https://doi.org/10.3390/electronics8050554

Tang, L., Zhang, Z., Zhao, Y., Feng, T., Wong, L., Garg, H., 2019: *A Comparison of WiFi-based Indoor Positioning Methods*. https://doi.org/10.1109/ICSPCS47537.2019.9008751

Yang, Y., Yang, H., Meng, F., 2025: A Bluetooth Indoor Positioning System Based on Deep Learning with RSSI and AoA. *Sensors*, *25*(9), 2834.

Zholamanov, B., Saymbetov, A., Nurgaliyev, M., Bolatbek, A., Dosymbetova, G., Kuttybay, N., Orynbassar, S., Kapparova, A., Koshkarbay, N., Beyca, Ö. F., 2025: RSSI Fingerprint-Based Indoor Localization Solutions Using Machine Learning Algorithms: A Comprehensive Review. *Smart Cities*, *8*(5), 153.

Zhu, X., Feng, Y., 2013: RSSI-based Algorithm for Indoor Localization. *Communications and Network*, *05*, 37–42. https://doi.org/10.4236/cn.2013.52B007