RESEARCH ARTICLE

DeepRSSI: Generative Model for Fingerprint-Based Localization

NAMKYUNG YOON [ID][1], WOOYONG JUNG [ID][1], AND HWANGNAM KIM [ID][1,2], (Member, IEEE)

[1]School of Electrical Engineering, Korea University, Seoul 02841, South Korea
[2]kuhnix Inc., Seoul 02841, South Korea

Corresponding author: Hwangnam Kim (hnkim@korea.ac.kr)

ABSTRACT In this paper, we present an innovative methodology for generating virtual received signal strength indicator (RSSI) fingerprint maps to improve indoor localization systems and wireless communication systems using RSSI. Focusing on the challenge of extensive labor and time required in traditional data collection, we propose a generative model that combines customized attention mechanism with a conditional variational autoencoder (cVAE), leveraging datasets compiled from direct measurements of RSSI values from different access points in a real-world indoor environment. Our model uniquely synthesizes high-quality virtual RSSI maps, significantly reducing the need for extensive physical data collection while enhancing the accuracy and efficiency of indoor positioning systems. By integrating measured data with innovative data generation techniques, our approach offers a novel solution to indoor localization challenges. In addition, this model can augment high-quality synthetic data for indoor wireless signals to expand the volume of available data. We quantitatively demonstrate the effectiveness of our model, showing an average improvement of over 40% in Euclidean distance errors across several machine learning algorithms compared to existing methods. Our experiments validate that the virtual RSSI fingerprint map yields accurate position estimates, with performance enhancements observed in algorithms that confirm the utility in real-world scenarios. The contribution of our research improves indoor localization systems by improving indoor positioning accuracy and addresses the limitations of traditional fingerprinting methods, setting the stage for future innovations in wireless communication.

INDEX TERMS Deep learning, fingerprint map, generative model, indoor localization.

I. INTRODUCTION

Wireless connectivity has become the backbone of modern digital life, with WiFi networks prevalent in various indoor environments [1]. Control, management, and performance optimization of these networks often utilize the received signal strength indicator (RSSI), a measurement that evaluates the power level of WiFi receiver is receiving from an access point (AP). Despite the importance of RSSI data, accurate measurement and prediction of RSSI data has long been a complex problem due to the irregular nature of indoor radio waves, which are primarily affected by factors such as

interference, obstructions, and multipath fading. Therefore, various studies are being conducted to obtain RSSI data accurately limited by various factors [2]. To address this, numerous studies are investigating methods to accurately capture RSSI data limited by these factors [3].

Our research is particularly relevant in the context of indoor localization, where high performance and accuracy are important [4], [5]. Fingerprint-based localization is one of the widely used methods for indoor localization, which records the RSSI of various APs at all locations. However, this method is typically labor intensive and time-consuming and requires extensive site surveys. A key challenge is to reduce the time and effort required for in situ data collection and provide an alternative to generating virtual fingerprint

The associate editor coordinating the review of this manuscript and approving it for publication was Prakasam Periasamy [ID].

maps without comprehensive in situ data. Furthermore, the performance of RSSI positioning systems critically relies on access to accurate RSSI data embedded within fingerprint maps. While deep learning algorithms readily capitalize on these data to improve positioning accuracy, a significant bottleneck lies in data acquisition; securing sufficient volumes of high-quality training data remains a persistent challenge [6]. This underscores the urgent need for robust RSSI data augmentation methods. Existing studies in this domain focus mainly on replicating existing data for augmentation purposes [7], [8], [9], [10], [11]. However, these methodologies often fall short in adequately capturing the intricate variability and multifaceted nature of RSSI data, leading to constraints in both data exploitation and system adaptability.

On the other hand, the generative model is a cornerstone in machine learning research, especially when working with limited or imbalanced datasets [12], [13]. It is a technique for creating new data instances that mimic the underlying characteristics of the original data, thereby enabling models to learn more robust and generalizable patterns. In the context of WiFi RSSI fingerprint map, generation can help alleviate several issues of certain signal strength levels or locations such as overfitting, and imbalance in data from different APs [14]. Moreover, the model is designed to dynamically account for indoor environmental variables such as obstructions and human movement, ensuring robust performance in real-world scenarios.

In this study, we propose DeepRSSI, an innovative model for generating a WiFi RSSI fingerprint map to improve fingerprint-based localization. Through this, generating a virtual fingerprint map reduces the need for laborious field work in fingerprint mapping, improving the efficiency and accuracy of indoor localization. Addressing these challenges underscores the vital importance and urgency of our research.

Our approach focuses on the innovative concept of virtual WiFi RSSI generation, a method that not only improves the accuracy of indoor localization, but also improves the performance of existing artificial intelligence-based RSSI algorithms through data generation. RSSI positioning techniques, which has the disadvantage of requiring a lot of existing resources, can be improved as a more influential tool by effectively performing indoor positioning through the data generated by our model. This dual advantage, which can also enhance existing applications and enhance the required data by generating precise virtual maps by properly learning the WiFi RSSI data, suggests significant advances in the field of WiFi-based indoor localization and network optimization.

Contributions of this paper: (1) Novel generative model. In this paper, we propose a high performance generative model based on a novel attention mechanism as sequential gate self-attention for DeepRSSI. The sequential gate self-attention mechanism enhances the ability of the model to focus on the most informative parts of the data sequence and to incorporate time-dependent positional information in the learning process, thereby achieving

superior results. In addition, to ensure the reliable learning of the model by the sequential gate self-attention and CNN layers, we utilize a loss function that incorporates a reparameterization trick based on Gaussian distribution. Our approach exploits the intrinsic temporal structure in RSSI data and the expressive power of conditional variational autoencoder (cVAE) to generate high-quality synthetic RSSI data.

Contributions of this paper: (2) Empirical experimental data. Another novel aspect of the research is our use of self-measured RSSI data. Utilizing the unique self-measured RSSI dataset, we offer a comprehensive evaluation of our model, including visual comparisons of RSSI contour maps, quantitative analysis of deviations, and simple discrimination models. The utilization of self-measured data not only adds authenticity to our research but also increases the applicability of our model, ensuring it caters to real-world scenarios. To validate the effectiveness of the proposed model, this paper presents a comprehensive evaluation study. This study includes visual comparisons of RSSI contour maps and quantitative analysis of deviations to evaluate the improved performance of data augmentation. This comprehensive evaluation method aims to provide a clear view of how our proposed generative model can enhance the quality and utility of RSSI data in indoor wireless networks.

Contributions of this paper: (3) Enhanced Indoor Localization. The proposed DeepRSSI model can reduce the time and effort spent on field investigations by generating virtual fingerprint maps without thorough field investigations, while maintaining or reducing localization errors. An innovative DeepRSSI method that conditionally generates a virtual RSSI fingerprint map not only reduces the need for extensive site investigations but also provides the possibility to reduce localization errors. Accordingly, our study represents a practical leap forward that makes the process of indoor localization-related tasks more streamlined and more efficient and accurate.

The overall goal of this study is to improve the robustness and abundance of RSSI data and optimize fingerprint-based localization performance through virtual fingerprint map.

The remaining part of the paper is organized as follows. Section II introduces detailed descriptions of the characteristics of RSSI data and various data generation techniques. Section III is the structure of the generative model based artificial intelligence for RSSI data, and Section IV is the experiment and performance evaluation result of DeepRSSI. Section V is the conclusion of this paper.

II. PRELIMINARY
A. RECEIVED SIGNAL STRENGTH INDICATOR
RSSI measures the power level of a signal received from a wireless network and functions as a relative measure of signal strength from a specific AP. RSSI data has diverse applications, such as calculating the location of individuals or objects, and even in the creation of routing

TABLE 1. Comparison of generative models in fingerprint-based localization.

Study	Methodology	Objective	Dataset	Outcome
Seong et al. [15]	Dual radio mapping with selective unsupervised AE and GAN	Reduce time-cost for Wi-Fi signal acquisition	Own dataset	Generate 88.59% accuracy signal map with 43.3% small size
Njima et al. [16]	GAN for fake fingerprints data augmentation	Enhance indoor localization with limited data	UJIndoorLoc [17]	Improving localization accuracy by 9.66%
Serreli et al. [18]	Fingerprint-based Positioning with GAN for LTE Signals	Improve accuracy with synthetic RSSI models	Own dataset	Accuracy in indoor (0.4m) and outdoor (10m)
Ai et al. [19]	Anomaly detection in fingerprinting with RAD-GAN	Ensure reliability of radio map	Own dataset + UJIndoorLoc [17]	Average 40% improvement in anomaly rate and AUC 0.96 anomaly detection
Suroso et al. [20]	VAE for RSSI synthetic data generation	Database augmentation for localization	Own dataset	Low average error in synthetic RSSI (-3dBm)
Njima et al. [9]	Selective GAN for RSSI data augmentation	Improve indoor location accuracy with synthetic data	UJIndoorLoc [17]	Localization accuracy improvement of 21.69%
Suroso et al. [21]	GAN for synthetic fingerprint database	Address offline small fingerprint database challenges	Own dataset	Improve localization 60% on average with synthetic RSSI data
DeepRSSI (Our Work)	**cVAE with a novel attention mechanism**	**Conditionally generate new maps for improved indoor localization**	**Own dataset**	**Average improvement over 40% in Euclidean distance errors**

tables [22]. Techniques utilizing Wi-Fi and bluetooth low energy (BLE) for indoor location measurement are precise but infrastructure-intensive and consume a lot of resources such as cost and time. Previous studies have explored the use and characteristics of Wi-Fi RSSI data to enhance wireless communication [23], [24].

RSSI also introduces variability in signals due to several factors, such as physical layout and fading of the environment, which can lead to inherent variability. Therefore, our objective is to study a model that learns with identifiable patterns attributed to the physical and structural characteristics of indoor environments based on the periodic characteristics of the RSSI data studied previously [25].

Furthermore, our study differentiates itself from these previous studies because it not only analyzes RSSI data, but also proposes a new model for the generation of this data, which is designed to dynamically adapt to various indoor environmental conditions such as obstacles and human movement

B. GENERATIVE MODELS
Generative models such as the following are widely used to enrich datasets, especially when the amount of available data is limited. Autoencoder (AE) is primarily used for data compression and noise reduction [26]. They consist of an encoder that maps input to a lower-dimensional latent space and a decoder that attempts to reconstruct the original data from the latent representation. While AE can learn to replicate the training data accurately, they may struggle to generate novel samples that are not directly based on the training data. Variational auto encoder (VAE) is a class of generative models that have gained considerable attention due to their ability to model complex, high-dimensional data distributions [27]. VAE maintains the characteristics of there

AE, but increases the diversity of the models output by adding a random variable to the latent representation. In this process, more colorful data can be generated while being close to the original data. Shorten and Khoshgoftaar [28] presented a comprehensive overview of data generation techniques in various domains. In general, GAN is widely used in various fields and applications such as images [29]. However, a recent study has shown that the effectiveness of VAE has improved with data generation to improve the performance of classification models [30].

In order to generate sequential table RSSI data that have inherently stochastic properties, we chose to advance the architecture of VAE due to their efficiency in generating various sample arrays suitable for this task. This selection is particularly relevant given the nature of RSSI data that require a subtle understanding of stochastic behavior for accurate generation. Although GAN excels at capturing and replicating high-dimensional data distributions, it often faces challenges in terms of training stability and mode collapse. On the other hand, the VAE-based generative model we target seeks to more effectively encode the spatiotemporal dependencies of the data through the latent space and learn reliably for the ability to create structured latent spaces that are important to RSSI data.

Therefore, we extend the capabilities specifically for generating RSSI data based on the unique strengths of the VAE. The enhanced features of DeepRSSI based on the VAEs architecture are designed to better capture the temporal and spatial dependencies inherent in RSSI data, which are often overlooked in conventional generative models. Therefore, these improved VAE architectures not only generate various RSSI data samples but also adequately learn the actual signal changes these samples are seen in the indoor environment.
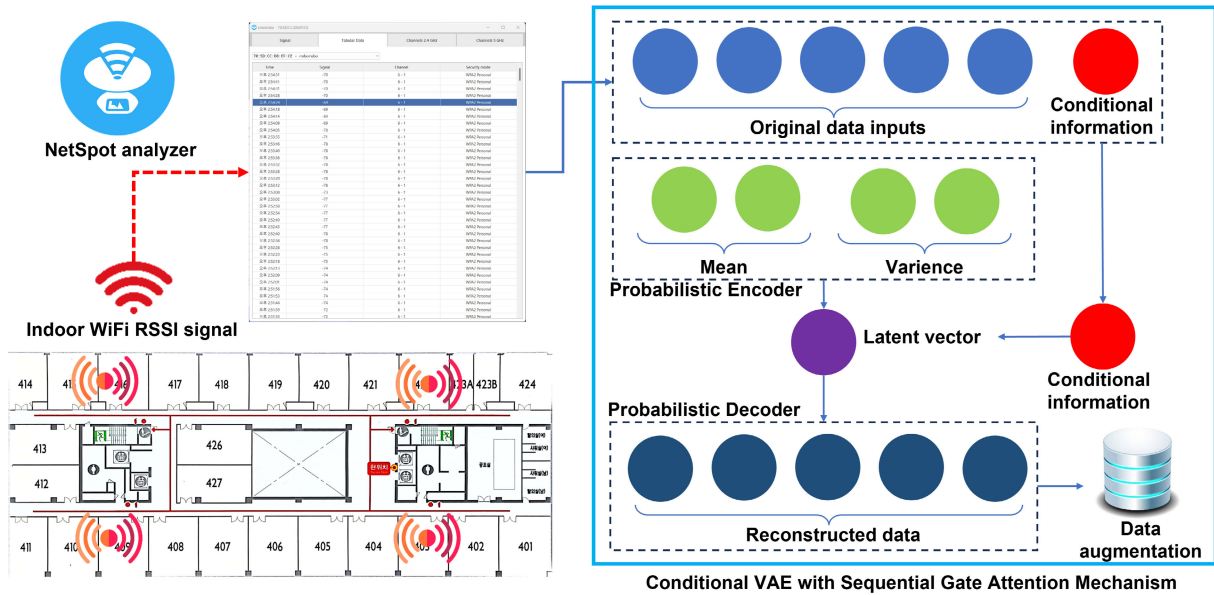
FIGURE 1. The procedure of generating indoor WiFi RSSI fingerprint map data through artificial intelligence model.

(a) Heatmap at AP 'A'

(b) Heatmap at AP 'B'

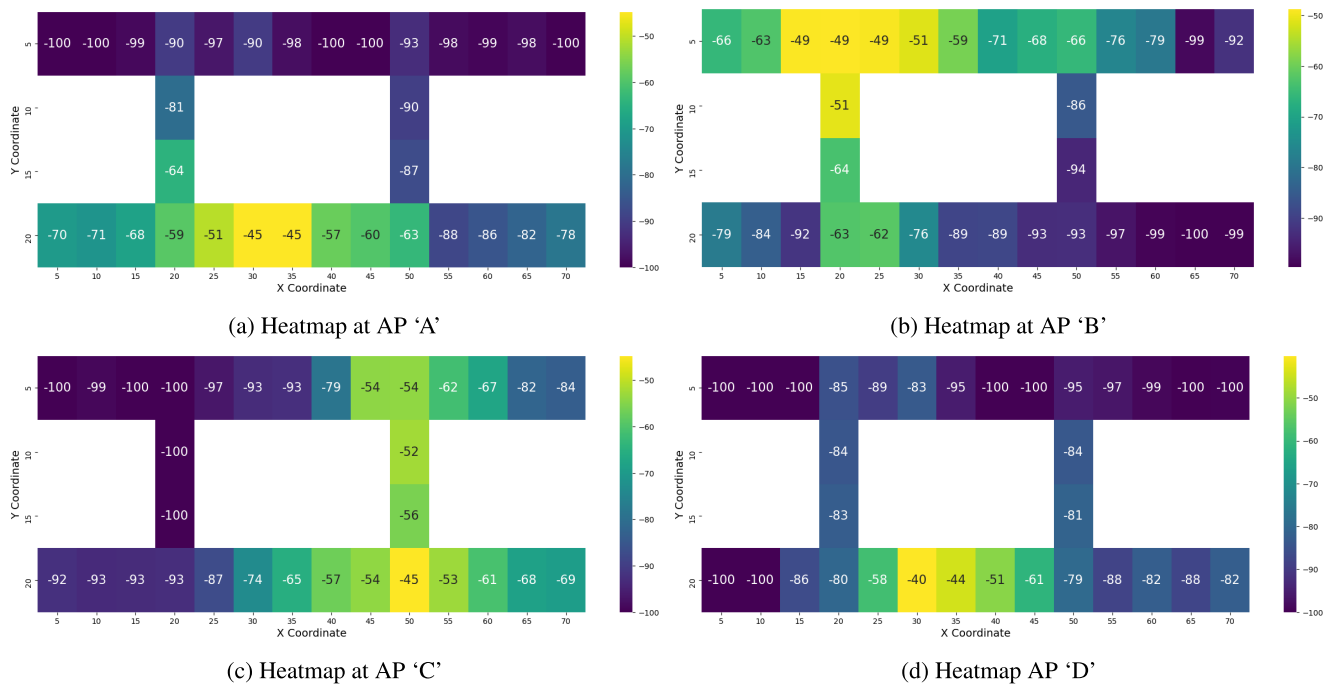(c) Heatmap at AP 'C'

(d) Heatmap AP 'D'

FIGURE 2. Fig. 2a, Fig. 2b, Fig. 2c, and Fig. 2d are the heat maps of the data measured on the fourth floor.

In following section, we design a system to demonstrate the adaptability and effectiveness of our DeepRSSI based on the characteristics of RSSI data.

C. SELF ATTENTION MECHANISM

The attention mechanism is a technique for efficient calculation by selectively focusing on specific features or areas of input data [31]. Attention mechanisms are used in various ways such as additive attention, dots attention,

and self-attention [32], [33]. The self-attention mechanism is a specific type of attention mechanism for processing sequential data, which has been increasingly used in various applications such as natural language processing or sequence-to-sequence tasks [33]. It has the ability to model dependencies regardless of their distance in the sequence, capturing long-term dependencies, and contributing to more accurate representation of sequences. This mechanism calculates the relevance of each element in a sequence to all other elements,
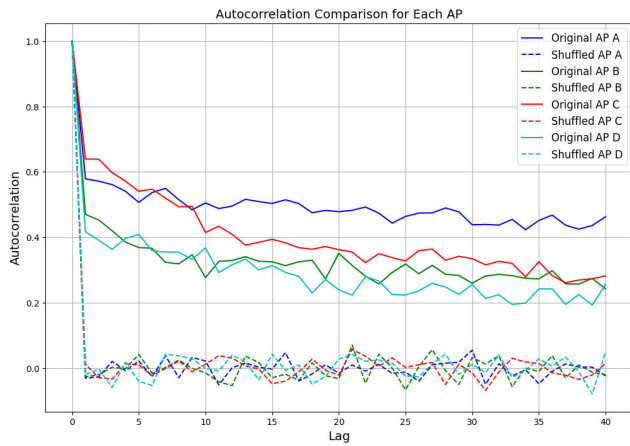
FIGURE 3. Compare autocorrelation between original RSSI data and shuffled RSSI data for each AP.

allowing the model to focus on the most relevant parts during data processing [34].

The self-attention mechanism that sorts and weights the importance of various functions within a sequence becomes a key innovation in our sequential gate self-attention mechanism. Our study builds on previous research in these areas but significantly extends it by proposing a unique model for RSSI data generation.

D. NOVELTY OF DEEPRSSI

Advances in indoor localization through RSSI-based fingerprinting have focused mainly on improving database quality and coverage. Studies by Seong and Seo [15] and Serreli et al. [18] in Table 1 have used generative models such as AE and GAN to augment fingerprint data, and aimed to reduce manual work required for data collection and improve positioning accuracy through optimized synthetic RSSI models. These studies mainly augment existing datasets without adaptability to specific requirements of various indoor scenario and characteristics of radio waves or WiFi signal.

In contrast, DeepRSSI introduces a subtle approach using a cVAE combined with a novel sequential gate self-attention mechanism. The method goes beyond simple data augmentation to conditional generation for a virtual RSSI map. It is a significant advance from the general data augmentation of existing studies, providing an efficient solution to adapt to the unique characteristics of different environments and generating maps that reflect the complex signal propagation dynamics within the indoor space. This feature distinguishes our approach from others, such as the augmentation-oriented tasks such as Suroso et al. [20], [21] and Njima et al. [16], [9], or in anomaly detection-oriented RAD-GAN such as Ai et al. [19] in Table 1.

Therefore, DeepRSSI is a strategic advancement that goes beyond the static generation approach seen in previous studies in Table 1 for RSSI data, and the conditional

generation of virtual maps to address the labor intensive deployment of fingerprint databases is a distinct contribution from existing studies.

III. DESIGN PRINCIPLE AND ARCHITECTURE

In this section, we detail the design principles and architecture of his approach to create a virtual fingerprint map, as shown in Fig. 1. We present DeepRSSI, an architecture of cVAE with various deep learning techniques that conditionally control floor data to generate RSSI data. In addition, we propose a novel attention mechanism based on the self-attention that allows the original RSSI data to be effectively generated using cVAE as shown in Fig. 4.

1) ANALYTICAL APPROACH FOR RSSI DATA

We designed the model to generate an effective fingerprint map for accurate indoor localization while preserving the spatial properties of the environment. The model aims to capture the underlying distribution of the strength and sequence data for RSSI signals associated with each access point (AP), while taking into account the location and floor information. We represent the importance of RSSI data sequences at APs (A, B, C, D), as shown in Fig. 3. This visualization compares the original RSSI signal sequence with the shuffled RSSI signal sequence for each AP, demonstrating the importance of sequence order in RSSI data analysis. The *Lag* in Fig.3 signifies the temporal delay between consecutive RSSI readings. In the unshuffled original RSSI data, this delay is characterized by a high autocorrelation, indicating a strong sequential dependency. Conversely, when the RSSI sequence is randomized, a marked decrease in autocorrelation is observed across all APs. This underscores the necessity of retaining the sequential integrity of RSSI data for the precise analysis and learning of artificial intelligence models.

We also clearly indicate the spatial variation of RSSI data in the indoor environment through a heat map showing the distribution of the average signal intensity of each AP, as shown in Fig. 2. These heat maps show the spatial variability and signal distribution patterns implied by each AP in RSSI data and highlight the need for models that accurately interpret and utilize these data.

Overall, Figs. 2 and 3, measured at each AP, highlight that both temporal and spatial aspects must be considered for proper learning of RSSI data.

A. SEQUENTIAL GATE SELF-ATTENTION MECHANISM

One of the major components of our DeepRSSI is the sequential gate self-attention mechanism. This mechanism builds on the principle of the transformer based self-attention mechanism [33]. However, we introduce several modifications that make it particularly suitable for processing sequential RSSI data. Without this mechanism, learning from RSSI data proves challenging and unstable due to the inherent complexities about WiFi RSSI data. The proposed mechanism enables superior learning of WiFi RSSI patterns
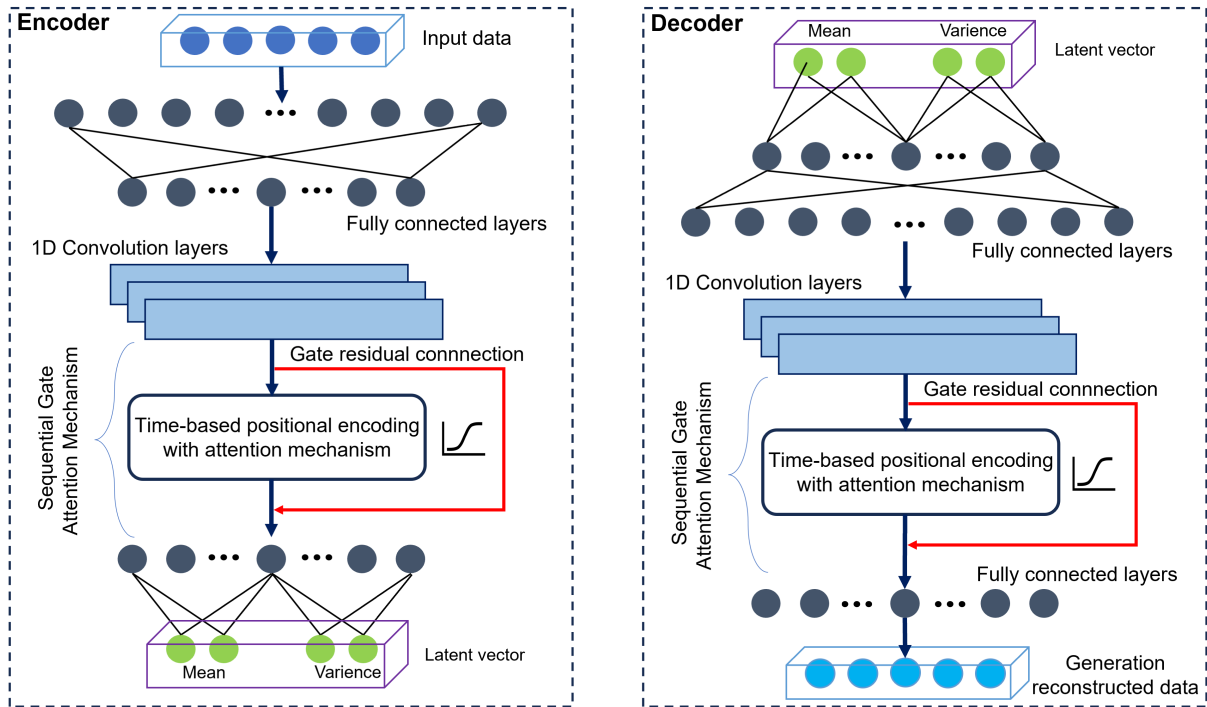
FIGURE 4. The overall structure of encoder and decoder layers of DeepRSSI model.

by effectively considering complexity and has been proven in real-world implementations.

The objective of self-attention is to calculate an attention score for each element in the sequence, reflecting the relevance of that element in the context of the other elements [33]. We introduce the sequential gate self-attention mechanism to consider the relationships among different RSSI measurements, thereby enhancing the learning ability of model to simulate realistic WiFi RSSI patterns.

Our attention mechanism assigns different weights to different vectors in an input WiFi RSSI sequence $x_1, x_2, \ldots, x_n$. In addition, this calculates queries (Q), keys (K), and values (V) for each element. These are calculated by applying linear transformations to the input sequence with time-based positional encoding (TPE), as shown in:

$$inputs = (x_1, x_2, \ldots, x_n) + TPE, \quad (1)$$

$$Q = inputs \cdot W_q, \quad (2)$$

$$K = inputs \cdot W_k, \quad (3)$$

$$V = inputs \cdot W_v. \quad (4)$$

In these equations, W_q, W_k, and W_v are weight matrices for the query, key, and value computations, respectively.

The attention mechanism in our model works by calculating attention scores and attention weights. The attention scores are determined based on the dot product of the query and key in the RSSI dataset, with a scaling factor to normalize the results. Attention weights are then computed using a softmax function applied to attention scores, ensuring a distribution between 0 and 1, which allows them to be

interpreted as probabilities. This process is fundamental in determining the relative importance of each element in the RSSI data sequence.

The attention scores are then calculated based on the dot product of the query and key for each pair of elements in RSSI dataset, scaled by the square root of these dimensions. The attention weights are then calculated by applying the softmax function to the attention scores. The softmax function ensures that the attention weights sum to 1 and are distributed between 0 and 1, which allows them to be interpreted as probabilities. Attention scores and attention weights are calculated as shown in Eqs. (5)-(6) below:

$$AttentionScores = \frac{QK^T}{\sqrt{d}} + \alpha(Q, K), \quad (5)$$

$$AttentionWeights = softmax(AttentionScores), \quad (6)$$

where d is the dimension of the query and key Q and K. K^T denotes the transpose of the matrix K, and the $\alpha(Q, K)$ function is designed to introduce an additional layer of contextual information into the attention mechanism by integrating the variances, which are statistical measures of the query and key. The function $\alpha(Q, K)$ is computed as the product of the variances of each vector Q and K. This computation not only enhances the attention scores but also embeds deeper insights into the data distribution, reflecting the inherent variability and dynamics of the RSSI data.

Finally, the output is calculated as a weighted sum of gated residual connection with value, where the weights are given

by the attention weights as shown in the following equation:

$$Output = inputs + Gate \cdot (AttentionWeights \cdot V). \quad (7)$$

This allows us to generate fake data that is more realistic and context-aware. According to Eqs. (1)-(7), our model not only incorporates temporal features through attention weights, but also encodes information by considering the physical distribution of AP signals within the input data, allowing the output to reflect both temporal and spatial characteristics of the RSSI environment. The following subsections describe details of these techniques.

We aim to improve the ability of the model to capture temporal dependencies and intricate spatial patterns within the RSSI data. Therefore, we introduce a special attention mechanism that combines two additional methods that are used to learn the properties of RSSI data well.

1) TIME-BASED POSITIONAL ENCODING

The original self-attention mechanism treats the input elements independently of their positions in the sequence [35]. While this approach is generally applicable, it may not fully capture the characteristics of RSSI data, which can exhibit structured patterns under certain conditions, as well as random fluctuations due to dynamic environmental variables. Our enhanced model is designed to account for these unique aspects of RSSI data. In the context of Wi-Fi signal strength measurements, the order of data points is vitally important. To improve existing limitations, we introduce a novel *TPE* using in Eqs. (1)-(4). The TPE consists of two main components: the positional encoding (*PE*) and the time embedding (*TE*).

Firstly, we construct the positional encoding, denoted as *PE* used in the Transformer for a given position *pos* and dimension *d* which can be $2i$ or $2i + 1$ [33]. The elements of this matrix are calculated by the following function, which incorporates a balance of sine and cosine functions:

$$PE_{(pos,2i)} = \sin\left(\frac{pos}{10000^{2i/d}}\right), \quad (8)$$

$$PE_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{2i/d}}\right). \quad (9)$$

This encoding is added to the input sequence before it is passed to our attention mechanism, thereby allowing the mechanism to take the sequence order into account.

Then, we introduce a new process of *TE* for *TPE* that further strengthens the ability of model to capture temporal patterns. The *TE* is created by applying a dot product of the weight matrix W_{time} with the input sequence and passing the result through a sinusoidal function. The use of a sinusoidal function provides a continuous and differentiable way to encode time information, which is a fundamental concept in signal processing that allows us to represent any periodic function as a sum of sinusoidal components of different frequencies. This is achieved by applying a linear transformation followed by a sinusoidal function to the input

sequence as follows:

$$TE = \sin(W_{\text{time}} \cdot inputs). \quad (10)$$

The trainable weight matrix W_{time} with Glorot uniform distribution is used to prevent vanishing and to learn important temporal features in the data during the training process [36].

Then, we combine both the *PE* and *TE* to create the *TPE* as follows:

$$TPE = PE + TE. \quad (11)$$

With respect to the input sequence, the resulting *TPE* is both location and time aware, allowing the sequential gate self-attention mechanism to better capture and exploit the temporal patterns and dependencies inherent in WiFi RSSI data.

Specifically, the encoding techniques we introduce allow our AI model to dynamically respond to changes in the environment within the indoor environment. These factors are known to cause complexity in RSSI data, and in this regard, the adaptability of our model represents a significant step forward in addressing this complexity.

2) GATED RESIDUAL CONNECTION

In DeepRSSI, we introduce a gated residual connection that modulates the contribution of each element in the sequence based on its importance, thereby selectively focusing on the crucial elements. In practice, residual connections, which can bypass the self-attention layer for elements that are not considered important, are often used to improve model performance [33], [37].

The purpose of our gated residual connection is to modulate the contribution of output based on its relevance to the final sequence representation. This introduces an additional level of adaptivity to the model, allowing it to focus its attention on the most informative parts of the input sequence. Therefore, our model selectively controls how much information to let through by using gated residual connection unlike traditional attention mechanisms that equally treat the output.

This gated residual connection can dynamically adjust the contributions from different parts of the sequence, allowing the model to focus on more relevant features and enhancing its ability to discern meaningful patterns in the RSSI data.

The gated residual connection operates through a learned linear transformation applied to the input sequence, followed by a sigmoid activation function as shown in Eq. (12):

$$Gate = \sigma(W_{\text{gate}} \cdot inputs), \quad (12)$$

where W_{gate} is a weight matrix that is learned during training. Through this, the gated residual connection empowers our model to focus more on the salient parts of the input sequence as shown in Eq. (7), selectively amplifying important features and attenuating less meaningful ones. This ability is crucial for dealing with RSSI data, which often contains temporal

---

Algorithm 1 Sequential Gate Self-Attention

1: **procedure** Sequential gate self-attention mechanism(*inputs*)
2:     **procedure** build(*self*, *input_shape*)
3:         Initialize attention weights W_q, W_k, W_v
4:         Initialize gate weights W_{gate}
5:         Define *position* as a range of input sequence
6:         Compute *div_term* using $2i/d$ on the *position* for *i* in range of dimensions
7:         Compute $PE_{(pos,2i)}$ and $PE_{(pos,2i+1)}$ using sine and cosine on the *div_term* and *position* to form *PE*
8:         Initialize time weight W_{time}
9:     **end procedure**
10:     $inputs{+}= PE$
11:     $TE = \sin(W_{time} \cdot inputs)$
12:     $inputs{+}= TE$
13:     $Q = inputs \cdot W_q$
14:     $K = inputs \cdot W_k$
15:     $V = inputs \cdot W_v$
16:     $AttentionScores = \frac{QK^T}{\sqrt{d}} + \alpha(Q, K)$
17:     $AttentionWeights = \mathrm{softmax}(AttentionScores)$
18:     $weighted\_inputs = AttentionWeights \cdot V$
19:     $gate = \sigma(W_{gate} \cdot inputs)$
20:     $output = inputs + gate \cdot weighted\_inputs$
21:     **return** *output*
22: **end procedure**

---

patterns with variable importance across different time steps. The gated residual connection allows the model to adaptively capture the most informative aspects of the RSSI data.

B. GENERATIVE MODEL BASED ON CVAE

In this section, we discuss the principles and architectural details of the DeepRSSI model designed to generate RSSI data with the aid of cVAE model with sequential gate self-attention mechanism.

In the cVAE model as shown in Fig. 4, we employ an encoder-decoder architecture to map the RSSI data into a latent space and subsequently generate new data instances as shown in Eqs. (13)-(14):

$$Encoder : q_\phi(z|x, c) = \mathcal{N}(z; \mu_\phi(x, c), \sigma_\phi(x, c)), \quad (13)$$

$$Decoder : p_\theta(x|z, c) = \mathcal{N}(x; \mu_\theta(z, c), \sigma_\theta(z, c)). \quad (14)$$

The encoder, given an input x and a condition c, produces a Gaussian distribution over the latent variable z. The mean (μ) and standard deviation (σ) of this distribution are functions of the input data and the condition. In contrast, the decoder generates new data instances from the provided latent representations and conditions. For a given latent representation z and condition c, the decoder defines a Gaussian distribution over the reconstructed data x. The mean (μ) and standard deviation (σ) of this distribution are functions of the latent representation and the condition.

In addition, we use 1-D convolutional neural network (CNN) in the cVAE model to effectively extract features of various signal strengths implied by RSSI data, as shown in Fig. 2, and put them in a shape suitable for the proposed sequential gate self attention mechanism as shown in Fig. 4. Given an input x, the encoder defines a distribution $q_\phi(z|x)$ over the latent variable z, while the decoder defines the conditional distribution $p_\theta(x|z)$. The objective function of our model, known as the evidence lower bound (ELBO) plays a crucial role [27]. It ensures the fidelity of the reconstructed data and maintains the regularity of the latent space. Specifically, we achieve efficient learning between data fidelity and latent regularity by optimizing model parameters to stabilize and optimize ELBO. The traditional cVAE loss function which combined with mean square error (MSE) and Kullback-Leibler (KL) divergence optimizes both the reconstruction and distribution of RSSI data as shown in Eq. (15):

$$Loss : \mathcal{L}(\theta, \phi; x, c) = \mathbb{E}z \sim q\phi(z|x, c)[\log p_\theta(x|z, c)] - D_{KL}(q_\phi(z|x, c)||p(z)). \quad (15)$$

Algorithm 2 Generative Model Base cVAE

1: **procedure** cVAE(*input*, *coordinates*)
2:     $enco\_node = concatenate(input, coordinates)$
3:     Apply Dense layers with leaky_relu to *enco_node*
4:     Reshape *enco_node* for 1D Convolution
5:     Apply 1D Convolution to *enco_node*
6:     Apply sequential gate self-attention to *enco_node*
7:     Flatten *enco_node*
8:     $z\_mean = \mathrm{Dense}(latent\_dim)(enco\_node)$
9:     $z\_log\_var = \mathrm{Dense}(latent\_dim)(enco\_node)$
10:     $z = sampling([z\_mean, z\_log\_var])$
11:     $deco\_node = concatenate(coordinates, z)$
12:     Apply Dense layers with leaky_relu to *deco_node*
13:     Reshape *deco_node* for 1D Convolution
14:     Apply 1D Convolution to *deco_node*
15:     Apply sequential gate self-attention to *deco_node*
16:     $prediction = \mathrm{Dense}(input\_num)(deco\_node)$
17:     **return** *prediction*
18: **end procedure**

In this case, we need to stabilize VAE model with strong performance to efficiently learn large amounts of RSSI data. Consequently, we present subtle modifications to traditional sampling functions that contribute to new improvements to the original re-parameterization trick by inserting small constants for numerical stability into the exponent of the exponential function as the following Eq. (16):

$$z = \mu + \exp\left(\frac{1}{2}\log\left(\sigma^2\right)\right) * \epsilon, \quad (16)$$

where z denotes the sampled latent variable, and μ and σ represent the mean and standard deviation of the latent variable z respectively. Traditionally, a sample z would be
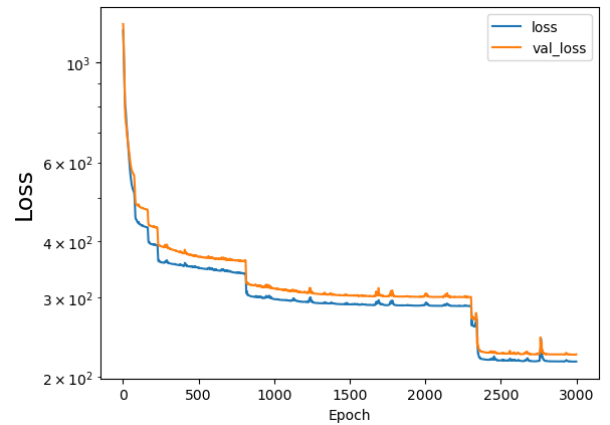
drawn from this distribution which is characterized by a mean μ and a standard deviation σ. However, this random sampling operation is non-differentiable and thus disrupts the backpropagation process integral to the training of neural networks. To circumvent this restriction, we use a standard normal distribution random variable ϵ. We apply a reparameterization trick that separates the stochastic part from the deterministic part of the network [27]. This ensures that z becomes a deterministic function of the parameters and ϵ, making the network fully differentiable from gradient-based optimization methods and serving as an effective means to stabilize the VAE model and efficiently learn from large amounts of RSSI data. The term $\exp\left(\frac{1}{2}\log\left(\sigma^2\right)\right) * \epsilon$ in the exponent is derived from the transformation of the Gaussian distribution, serving as an effective means to stabilize the VAE model and efficiently learn from large amounts of RSSI data. The log-variance is unbounded from above and below, which gives more flexibility during optimization. This contrasts with the variance itself, which is bounded below by zero. This small modification for numerical stability is a frequently used idea in the field of machine learning used according to the model and research purpose [27]. Therefore, we show through the practical implementation of Section IV that the trick contributed appropriately to the RSSI generative model. Through the sampling function that updates these processes, our cVAE model with CNN layer and Sequential gate self-attention mechanism is able to generate stable synthetic data that are similar to the training data and adherent to specific conditions, which in our case are the RSSI data. With the sampling function thus updated, the KL divergence computation is reformulated as shown in Eq. (17):

$$D_{KL} = -\frac{1}{2}\sum_{j=1}^{J}\left(1 + \log\left((\sigma_j)^2 + \epsilon\right) - (\mu_j)^2 - (\sigma_j)^2\right),$$
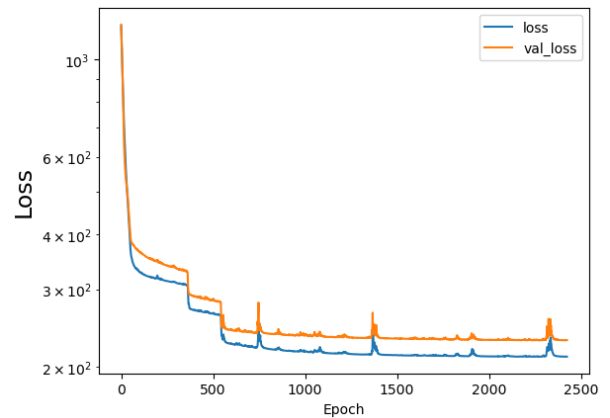
(17)

where D_{KL} is the KL divergence measuring the difference between the learned latent distribution $q_\phi(z|x)$ with ϕ being the parameters of the encoder and a prior distribution $p(z)$. Concurrently, $p_\theta(x|z)$ defined by the decoder parameter θ serves as a likelihood function that contextualizes the interaction between the encoder and decoder in the latent variable space z.

Through these processes, the proposed DeepRSSI model with sequential gate magnetism is able to generate stable synthetic data that are similar to the training data and adherent to specific conditions, which in our case are the RSSI data.
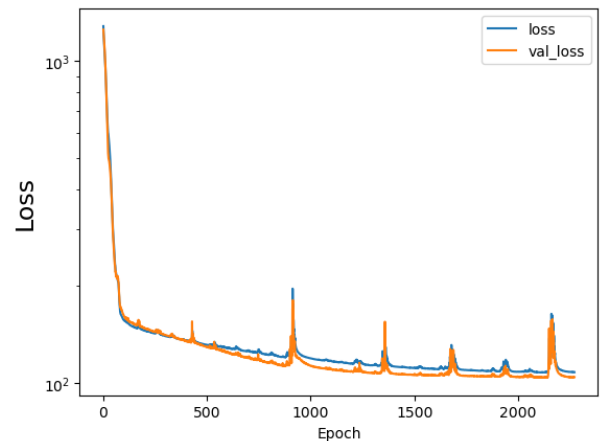
The DeepRSSI model, integrating a cVAE with our proposed attention mechanism, is meticulously designed for efficient computation, optimizing both time and space. The computational complexity of the sequential gate self-attention mechanism, denoted as $O(N^2 \cdot A)$ for input sequences of length N and attention dimension A, focuses computation on the most crucial parts of the data sequence N. This strategic focus not only explains the quadratic dependence on sequence length N but also the linear

(a) Loss of cVAE with only fully connected layers.

(b) Loss of cVAE with fully connected layers and 1D CNN layers.

(c) Loss of cVAE with fully connected layers, 1D CNN layers and sequential gate self-attention mechanism layer.

FIGURE 5. Comparison of cVAE in training process. The blue line represents training loss, and the orange line represents validation loss.

dependence on attention dimension A, underscoring the capability of model to efficiently process sequences. The inclusion of the self-attention mechanism results in a parameter increase from 31,056 to 31,840, translating to an approximate 2.52% increase in memory overhead. Despite
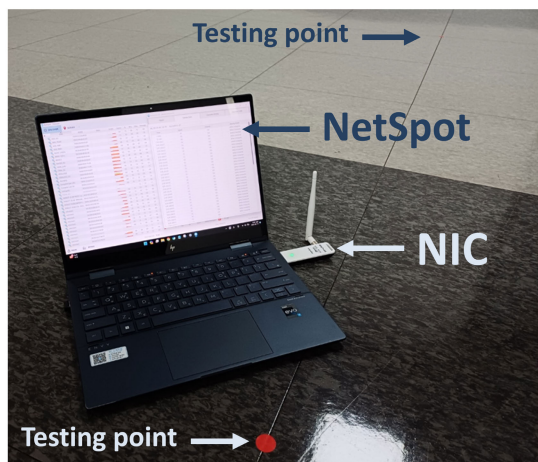
FIGURE 6. Experimental setup for RSSI data collection.

this nominal augmentation in resource requirement, the efficiency gained in model performance is significant. This enhancement is achieved by directing computational efforts towards the most informative segments of the data sequence, thus improving performance without incurring substantial additional memory costs.

IV. EXPERIMENTS

In this section, we evaluate the performance by comparing simulated and original data using quantitative and qualitative methods.[1] The DeepRSSI we present has the potential for a new paradigm shift that can leapfrog the existing RSSI fingerprint map technique, which is declining due to cost and time constraints, by conditionally generating similar simulation data rather than just augmenting it. Therefore, we demonstrate the effectiveness of various quantitative and qualitative evaluations and real-world indoor positioning of virtually generated fake data. We also demonstrate the performance of DeepRSSI using directly measured data from dynamic real-world environments and the necessity of using sequential gate self-attention mechanisms in DeepRSSI. The performance evaluation plan entails three core components: visual comparison of RSSI contour maps, analysis of quantitative deviations such as variance, and the application of simple discrimination models. These components provide a thorough and rigorous evaluation of the effectiveness and efficiency of the model in improving RSSI data quality, and thereby fingerprint map-based localization accuracy.

A. IMPLEMENTATION

We implement the sequential gate self attention mechanism proposed in Section III using Python code as shown in Algorithm 1 and use it in DeepRSSI model-based cVAE as shown in Algorithm 2 with parameters such as Table 2. In detail, we use 3000 epochs and the Adam optimization

---

[1]Data related to this paper are available upon request by contacting the author.

TABLE 2. Training parameters.

Parameter	Value
Epochs	3000
Optimization Function	Adam
Learning Rate	0.001
Loss Function	Mean Squared Error and KL Divergence
Data Split	90% Training, 10% Testing

function using a learning rate of 0.001, a mean squared error and KL divergence for loss function and use a callback function called early stopping to prevent overfitting. The loss function of a VAE consists of two main components. First, MSE is used to measure the similarity between the data learned by the model and the data generated through data augmentation. Second, the KL Divergence assesses how closely the latent vector produced by the encoder matches the normal distribution used for sampling. Traditional backpropagation is not feasible in VAE when sampling latent vector, so backpropagation is facilitated by multiplying the latent vector by the normal distribution. As training progresses, the mean and variance outputs from the encoder gradually converge towards the normal distribution through KL Divergence. This research use Intel i5-11500 CPU @ 4.60 GHz, 32 GB RAM, Nvidia GeForce RTX 3060ti for training and data generation. To evaluate the performance of model effectively, we allocate 90% of the data for training purposes to ensure a comprehensive learning process and reserve the remaining 10% for testing so that the generalization capabilities of model can be evaluated unbiased.

B. EXPERIMENTS IN REAL-WORLD ENVIRONMENT

1) DATA ACQUISITION

We conduct a wireless site survey using NetSpot on the Windows 11 Laptop with a network interface card (NIC) [38]. We conducted the experiments in a controlled indoor environment with a known layout and designated APs as shown in Fig. 1. This allowed us to directly capture real-time RSSI data from the surrounding Wi-Fi network. We also use data measured in an indoor environment without controlling for obstacles such as objects, human movement effects, and other variables that may affect the proposed model.

The dataset was obtained from a right-angled, densely populated building with numerous wireless networks, which represents common indoor environments. This enhances the reliability and applicability of our models and experimental outcomes. To ensure accurate and reliable measurements, we carefully calibrated the NIC and positioned the laptop at various locations within the indoor environment.

Figure 6 illustrates the experimental setup used in the RSSI data collection process. Figure 6 included a laptop with a NIC for recording RSSI values, with reference points strategically placed within the interior environment. Each red dot on the floor marks to a grid spaced 5*m* apart to display the reference locations where RSSI readings from the various

(a) Visual comparison for AP 'A' in floor 4

(b) Visual comparison for AP 'E' in floor 5

(c) Visual comparison for AP 'B' in floor 4

(d) Visual comparison for AP 'F' in floor 5

(e) Visual comparison for AP 'C' in floor 4

(f) Visual comparison for AP 'G' in floor 5

(g) Visual comparison for AP 'D' in floor 4

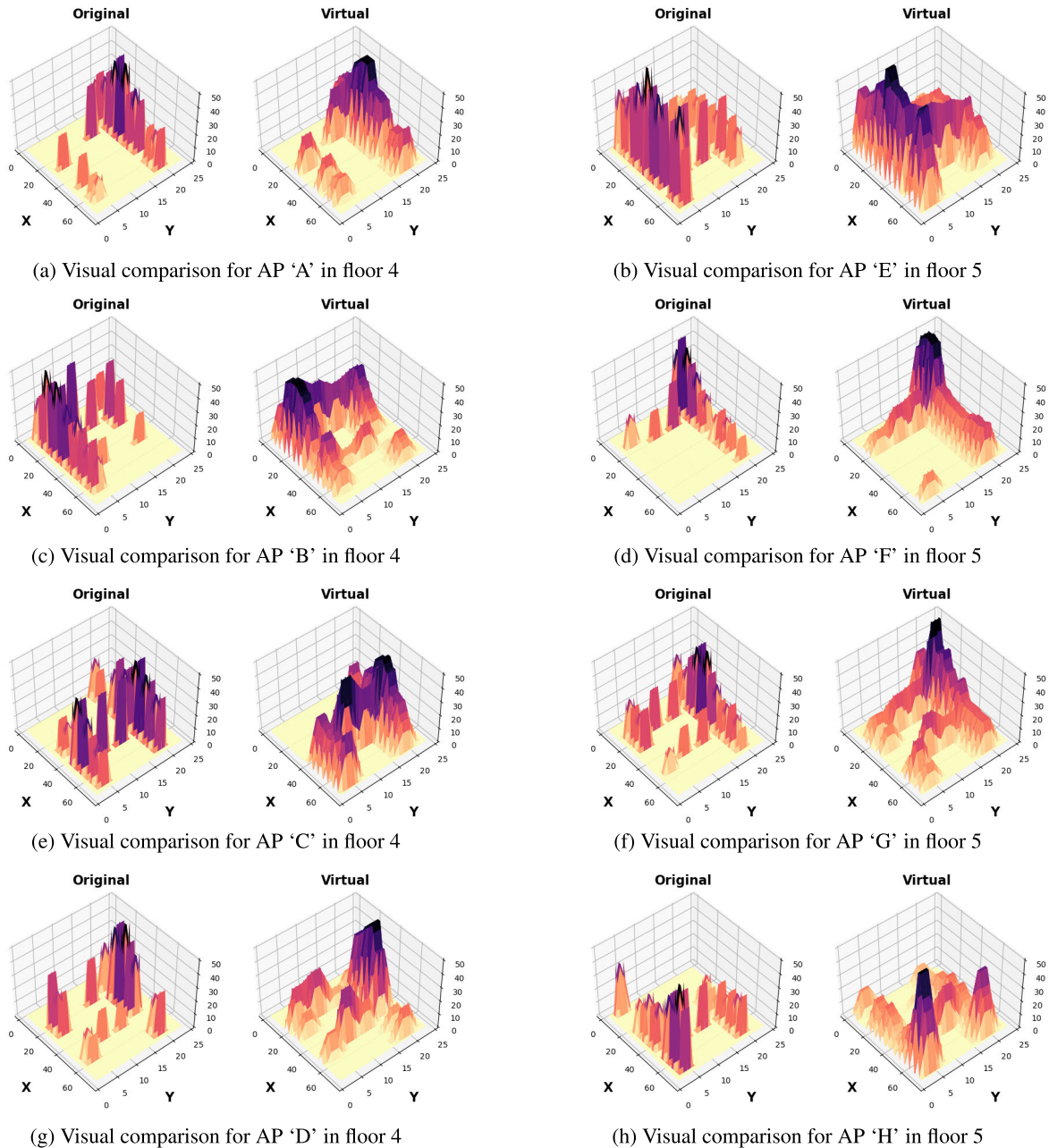(h) Visual comparison for AP 'H' in floor 5

FIGURE 7. Visual comparison between real RSSI data and virtual data at each AP.

APs are meticulously recorded. The selection of $5m$ intervals provides the basis for DeepRSSI to generate finer-interval virtual maps. This methodology demonstrates the ability of DeepRSSI to generate detailed indoor maps from sparse datasets, reducing the need for thorough data collection while maintaining high localization accuracy. Subsequently, further measurements are made with smaller intervals for testing for performance evaluation.

There are several advantages to measuring RSSI data directly using the Netspot program and a Windows 11 laptop equipped with a NIC. It enables us to capture real-time RSSI values from the indoor environment and mapping them to heatmap to help optimize AP deployment, providing a direct

and accurate representation of the Wi-Fi signal strengths at various locations [39]. This approach ensures the authenticity and reliability of the collected data, making it a valuable resource for training and evaluating the augmented model. In the following sections, we present the results and analysis based on this directly measured RSSI dataset, demonstrating the effectiveness and performance of our proposed generative model for enhancing indoor Wi-Fi RSSI data.

2) DATA STRUCTURE AND PREPROCESSING

The preparation of our experiments begins with the organization and processing of the RSSI dataset. This dataset, stored in comma-separated values (CSV) format, consists of

TABLE 3. PCC and SCC for each AP.

AP	PCC	SCC
4th Floor		
A	0.894	0.891
B	0.751	0.753
C	0.909	0.905
D	0.909	0.907
5th Floor		
E	0.792	0.848
F	0.764	0.682
G	0.878	0.862
H	0.886	0.864

columns indicating 'AP', 'sequence', 'signal', 'x', 'y', and 'z'. 'AP' represents one of the four APs from which the RSSI data ('sequence' and 'signal') are collected. We use datasets representing RSSI measurements on floors 4 and 5 of the building to learn the DeepRSSI model. Data on each floor consisted of 4481 measurements with data structures.

In this study, the scale of RSSI is measured from the minimum value of -100 dBm to the maximum value of 0 dBm. In the dataset, the minimum RSSI value that the device can measure is -99 dBm. Therefore, the value -100 dBm is used to indicate that there was no detectable signal.

The first stage of data preprocessing involves divided the complete dataset into distinct subsets, each corresponding to a specific AP, labeled as 'A', 'B', 'C', and 'D' in the fourth floor. In addition, the labels corresponding to the AP on the fifth floor are 'E', 'F', 'G', and 'H'. This division enables the handling of data pertaining to individual APs, facilitating a more granular analysis of the RSSI data. After the data preprocessing process, the resulting dataset has dimensions (256, 35). To utilize it efficiently, it was divided into training, validation, and test datasets at an 8:1:1 ratio. This division strategy was designed considering the overall distribution of the data to ensure optimal training and evaluation. Although the APs on each floor are different, they are similarly labeled to enable efficient execution of subsequent correlation tests, KS tests, and Q-Q plot tests.

3) PERFORMANCE EVALUATIONS
Prior to full-scale performance evaluation, we present a process by which the cVAE generative model stabilizes loss depending on the learning epoch, as shown in Fig. 5. Despite the complex and long structure of RSSI data, we can confirm that our designed CNN layer, sequential gate self-attention mechanism layer, and reparameterization trick enable fast and reliable learning. In addition, we use a convolutional filter that acts as an interpolation to create a smooth contour map for RSSI data. As can be seen visually, the contour map of RSSI generated by each AP shows that the original data and the composite data are very similar.

4) VISUAL COMPARISON
Visual comparison of RSSI contour maps serves as the qualitative analysis in the evaluation plan as shown in the Fig. 7. It allows for a graphic interpretation of the quality

of the fake data. This method is essential as it enables us to visually affirm the improvements in signal strength and coverage provided that each AP position is generated. Contour maps in visualization metric are generated to provide a spatial representation of the RSSI data across the various floors for each AP. In these maps, the color represents the RSSI signal strength, and the location of each point corresponds to the physical coordinates of the Wi-Fi signal across various layers for each AP. Visualization using these contours allows us to intuitively express Wi-Fi signal propagation patterns in the environment, thereby improving our understanding of the spatial distribution of signal strength in the environment.

5) QUANTITATIVE COMPARISON
The second aspect of the evaluation study concerns quantitative deviations, specifically variance. This analysis provides a statistical measure of the spread and dispersion of the fake data. A reduction in variance signifies less disparity in signal strength, indicating the success of model in providing consistent signal coverage across different indoor locations. We use Pearson and Spearman correlation coefficients to quantitatively compare the original data and the virtual data obtained through the generative model. The Pearson correlation coefficients quantify the linear correlation between the two datasets, while the Spearman correlation coefficients measure the strength and direction of the monotonic relationship between the two datasets. Both these metrics provide us with a comprehensive understanding of the correlation patterns of the original and virtual data.

Additionally, we present using the Kolmogorov - Smirnov (KS) test and the quantile - quantile (Q-Q) plot in a quantitative manner to verify the similarity between the original RSSI data and the fake data. The KS test is used to determine whether two distributions differ. However, there is a limitation that the KS test alone is not clear about the general difference because it is sensitive to each feature. Therefore, we further conduct Q-Q plot comparisons to conduct quantitative performance evaluations more comprehensively. The Q–Q plot compares the data with the quantile of the standard normal distribution to visualize how properties such as location and range of the distribution are similar or different in the distribution of the two data.

- Results of the correlation coefficients: As shown in Table 3, the Pearson correlation coefficient (PCC) shows 0.751 for the lowest AP 'B', and the remaining APs show a value approximately 0.9. The Spearman correlation coefficient (SCC), which measures the monotonic relationship between the datasets, also demonstrates a strong correlation, with the lowest being 0.753 for AP 'B', and the remaining APs yielding values around 0.9. Both PCC and SCC approach +1, indicating a strong positive correlation. Therefore, it can be interpreted that the virtual data exhibit a high average correlation of 0.8 or more, demonstrating a strong linear and monotonic relationship with the original data [44].
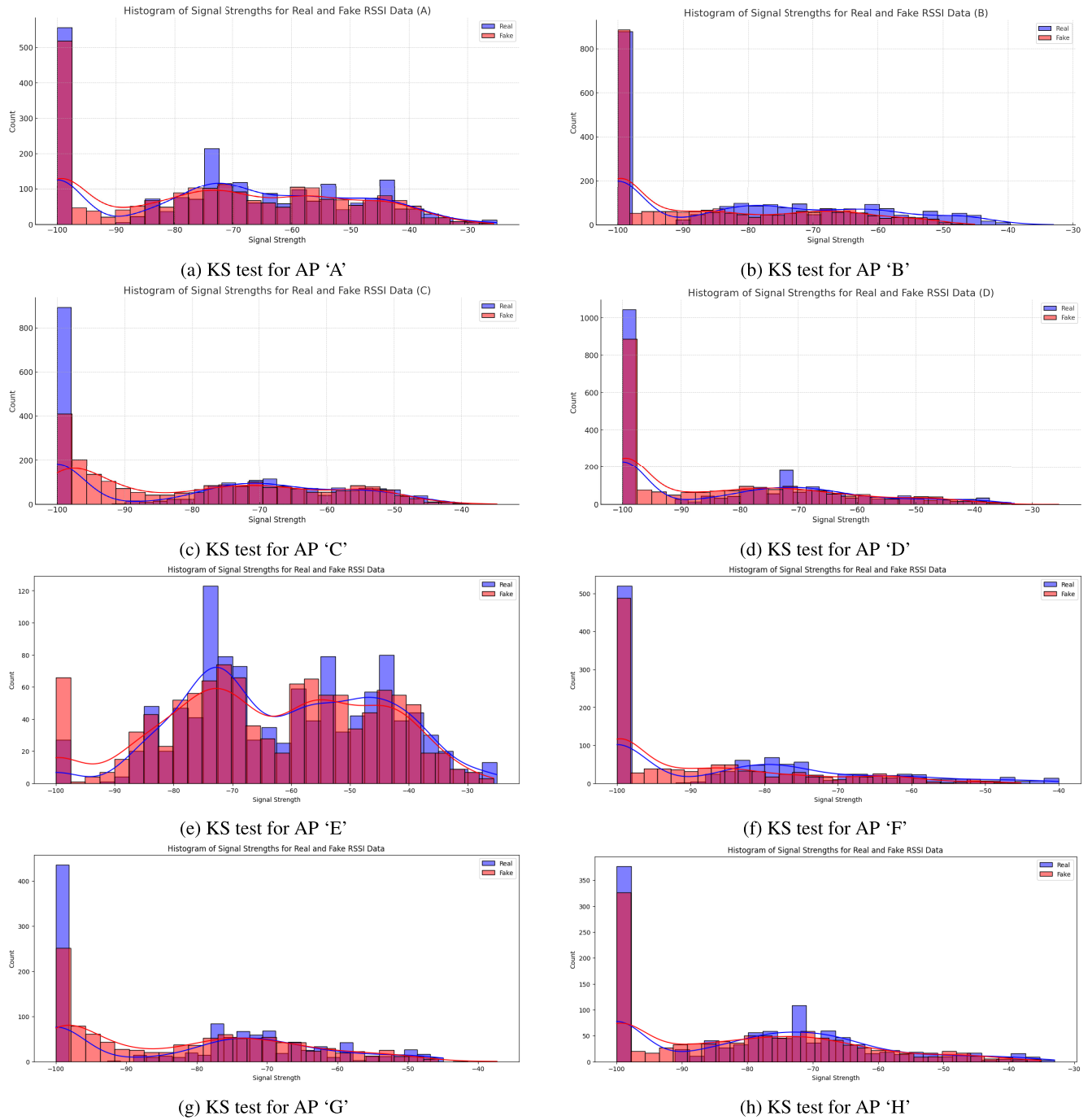
(a) KS test for AP 'A'

(b) KS test for AP 'B'

(c) KS test for AP 'C'

(d) KS test for AP 'D'

(e) KS test for AP 'E'

(f) KS test for AP 'F'

(g) KS test for AP 'G'

(h) KS test for AP 'H'

FIGURE 8. KS tests for each APs. Each histogram and trend line represent fake data in red and real data in blue.

- Results of KS test: In the Fig. 8, the KS test results across each of the APs demonstrate that there is a difference between the real and virtual data. However, this difference is relatively small, and the maximum difference in the KS statistic slightly exceeds 3%. In addition, each AP also mostly statistically indicates that the p-value is much smaller than the conventional significance level of 0.05. This suggests that although the distribution is not the same, the variation is minimized, indicating a

certain level of consistency between the real and virtual datasets. The overlay of trend lines on the histograms provides a visual affirmation of this similarity, as they closely trace each other, illustrating a consistent pattern across both real and virtual datasets.

- Result of quantile-quantile Plot: In the Fig. 9, a blue dot represents real data and a red dot represents virtual data. Gray lines are also fitted to the quantiles of real and virtual data, respectively. The Q-Q plot visually
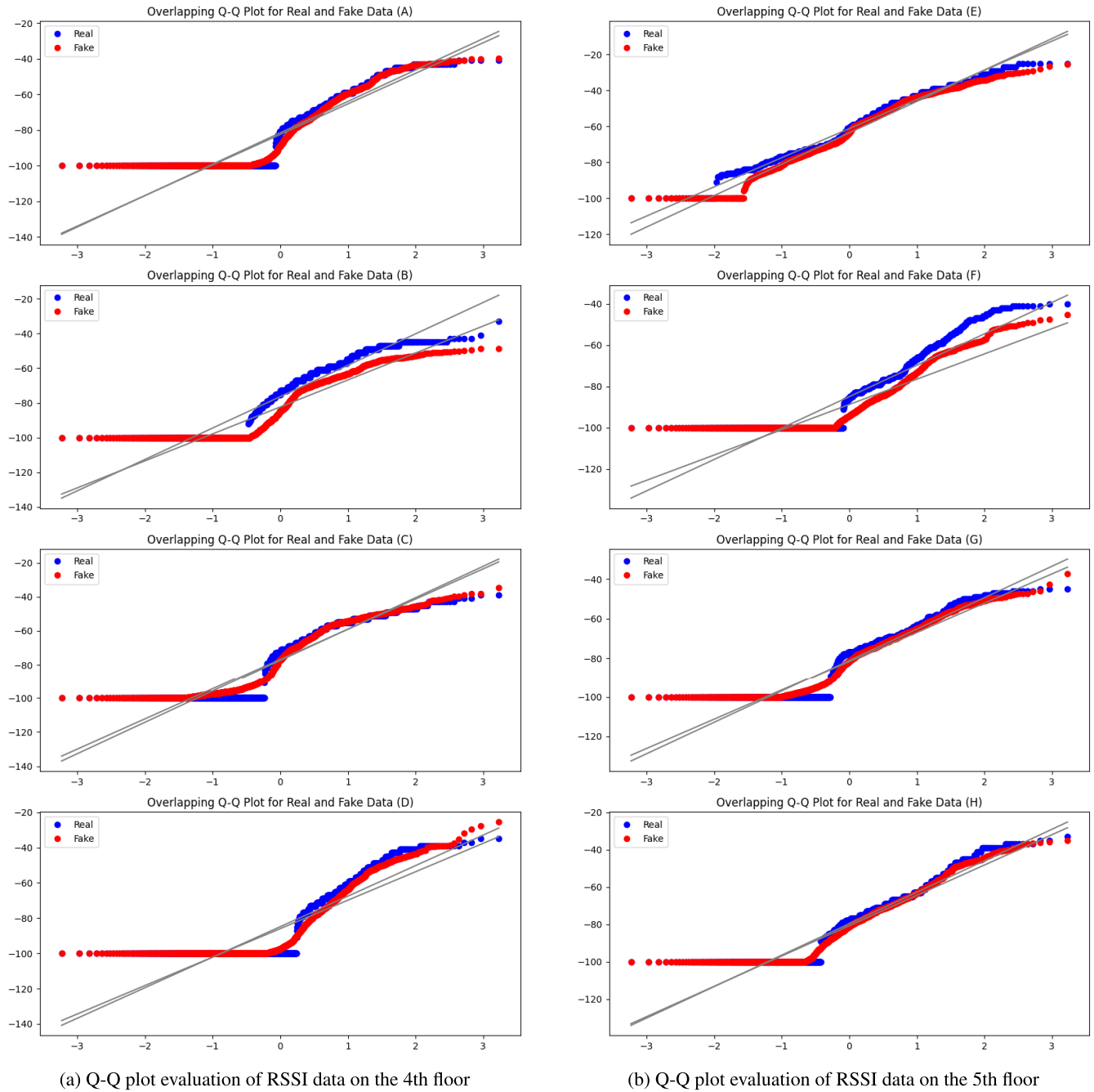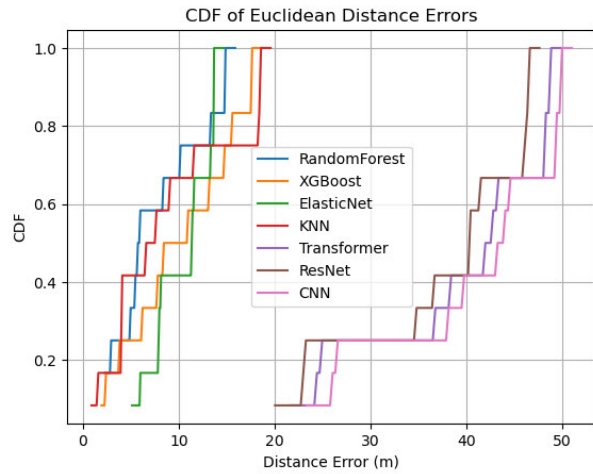
(a) Q-Q plot evaluation of RSSI data on the 4th floor

(b) Q-Q plot evaluation of RSSI data on the 5th floor

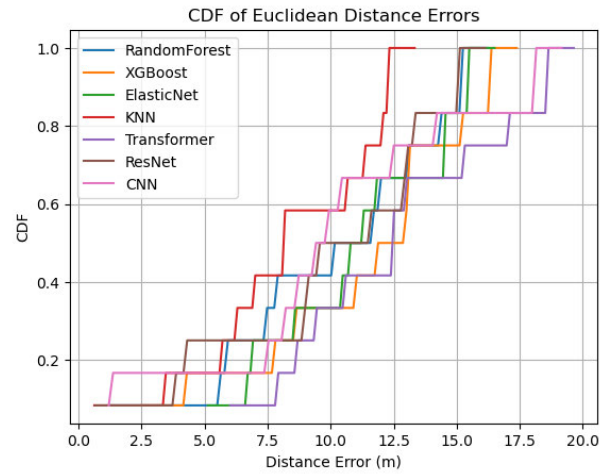**FIGURE 9.** Q-Q plot for each AP between real and virtual data.

compares the distribution of real and virtual data for each data set. The overlapping blue and red dots in the middle of the distribution means that the real and virtual data are somewhat similar in these regions. However, in the tail part of the distribution, there is a slight difference between real and virtual data. Nevertheless, these differences are tolerated in this study which prioritizes common patterns over outlier cases, and are demonstrated in subsequent indoor positioning experiments.

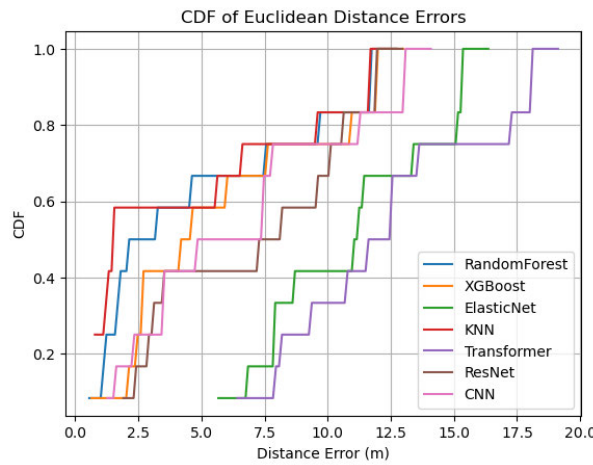### 6) INDOOR LOCALIZATION ERRORS

We compared the Euclidean distance error using existing commercialized machine learning algorithms and displayed the results in Fig. 10, which suggests the feasibility of the virtual fingerprint map. In addition, we strategically employ a 5m interval for real data collection, a distance that is sufficiently effective for indoor localization positioning [45]. The Fig. 10a is the indoor data of the fifth floor of the building measured at 5m intervals, and Fig. 10c is the virtual indoor data of the fifth floor generated at 1m intervals by the model
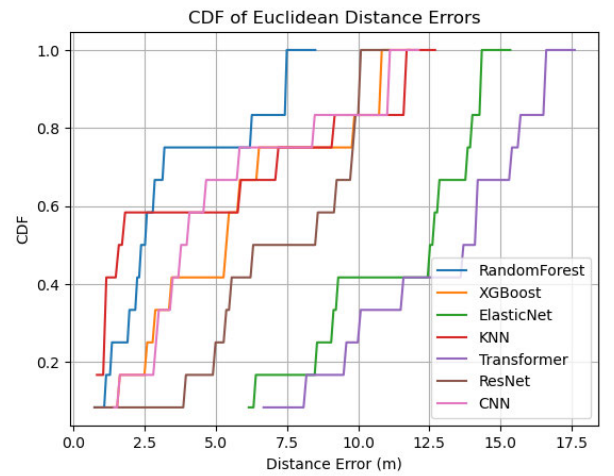
(a) Results of indoor measurement error using actual measured RSSI data on the 5th floor.

(b) Results of indoor measurement error using measured RSSI data of 5th floor at 1m interpolation.

(c) Results of indoor measurement error using virtual RSSI data of 5th floor at 1m intervals.

(d) Results of indoor measurement error using virtual RSSI data of 5th floor at 50cm intervals.

FIGURE 10. Comparison of CDF results that measured indoor localization errors using various learning algorithms based on virtual fingerprint map [40], [41], [42], [43].

trained with the data of the fourth floor. It is important to note that the training of model on the fourth-floor data does not directly influence the fifth-floor predictions. Our approach emphasizes the efficiency of generating a multi-layered virtual fingerprint map when the propagation environments of the floors are similar. This contribution suggests that, using data on a single layer, we can generate virtual fingerprint maps on multiple layers, which can significantly reduce the time and effort required for related tasks.

To evaluate indoor localization errors, we conduct two experimental studies using a variety of fundamental machine learning algorithms and deep learning architectures like Transformer, CNN, and ResNet model [41], [42], [43]. The Transformer model incorporates MultiHeadAttention layers and dense layers for processing, optimized with the Adam optimizer [41]. The CNN architecture utilizes sequential convolutional layers, focused on extracting spatial

hierarchies in data, also optimized with Adam [42]. The ResNet model employs convolutional and residual blocks, focusing on the preservation of features between layers [43]. These algorithms are currently used in indoor localization studies, and we experiment with each model with the default settings provided in the scikit-learn library [40]. As shown in Fig. 10, this study confirms that the results of our generative virtual dataset are better than the actual data, even if it is not measured in practice. In terms of quantitative terms, our virtual fingerprint dataset demonstrates the average performance improvement of Euclidean distance errors, which is improved over the actual measured fingerprint dataset, with RandomForest at 46.50%, XGBoost at 30.33%, KNN at 35.06%, Transformer at 64.65%, ResNet at 76.69%, and CNN at 76.22%. The performance of ElasticNet is similar to the actual measured fingerprint map, with a slight difference of about -4.5%. In addition, when we generate

smaller 50*cm* intervals fingerprint map with our generative model, as shown in Fig. 10d, the Euclidean distance improved by 33.38% for RandomForest, 8.94% for XGBoost, 6.21% for ElasticNet, 7.91% for Transformer, 14.37% for ResNet, and 14.00% for CNN compared to the virtual data generated with 1*m* intervals. Moreover, we compare the performance of the fingerprint map data of the virtual map generated at 1m intervals with the fingerprint map data made using linear interpolation at 1*m* intervals to prove the validity of this generative model. This study shows that the performance of each model when applied to the virtual map of as shown in Fig. 10c improved by 21.55% for RandomForest, 25.25% for XGBoost, 0.83% for ElasticNet, 4.72% for KNN, 2.62% for Transformer, 19.65% for ResNet and 26.45% for CNN compared to using linear interpolation as shown in Fig. 10b. The results indicate that our generative model has the potential to significantly reduce time and cost expenditures in the field of fingerprint-based indoor localization without compromising accuracy.

Overall, the KS test and Q-Q plot present a difference in empirical distribution between original data and virtual data, particularly in the tail regions. The generative model demonstrates excellent performance in the context of indoor localization. Our experiments on indoor localization errors show that the virtual dataset generated by our model not only reflects the main patterns found in real-world data, but outperforms them in terms of localization error. Furthermore, the performance of machine learning algorithms such as basic Random Forest, which are not tuned on our hypothetical fingerprinting map, shows competitive performance in a study that is actually positioned against that algorithm [5], [40].

In conclusion, while the virtual data are not an exact replica of the original data, they successfully simulate the key patterns and characteristics present in the original data, making them a significant and useful tool for tasks like indoor localization. The ability of the generation model to produce such high-quality virtual data highlights its potential utility in this domain. It allows us to achieve comparable accuracy while significantly reducing time and cost expenditures in site survey for fingerprint-based localization.

C. EFFECT OF VIRTUAL DATA FOR POSITIONING ACCURACY

We generate 32 real data samples and a total of 3612 virtual data samples to evaluate the impact of mixing virtual fingerprint data with real data on the accuracy of indoor positioning systems. From a dataset consisting purely of real data (0%) to one composed entirely of virtual data (100%), we experimented with different ratios of virtual to real data, including intermediate mixes of 25%, 50%, and 75%. The ratio of virtual data represents the proportion of virtual samples relative to the base number of 32 real samples. Specifically, for the 100% virtual data scenario, we utilized an equivalent number, 32 out of the 3612 virtual samples, to emphasize the robust generative capabilities of our model.

We evaluated indoor localization errors across multiple models using traditional machine learning algorithms and advanced deep learning architectures. The results shown in Table 4 highlight the potential benefits of virtual data integration in improving positioning precision across all models tested. The values in Table 4 represent the mean localization error in meters, which we determined by averaging the distances of the localization error in multiple iterations of the tests, each using a specific proportion of real and virtual data. The results show a clear trend: As the proportion of virtual data increases, the localization error tends to decrease, underscoring the efficacy of our data augmentation technique.

Our results show that positioning accuracy is significantly improved, especially for deep learning models such as Transformer, ResNet, and CNN, which outperforms existing algorithms with higher virtual data rates. For example, the accuracy of ResNet and CNN peaks at 100% virtual data rates, demonstrating their ability to effectively utilize synthetic fingerprints. Conversely, existing models show optimal performance at lower virtual data rates (25% and 50%), suggesting that balance is critical to maximizing the benefits of virtual data augmentation. This work highlights the importance of leveraging virtual fingerprints to improve indoor positioning systems, especially in scenarios where real data collection is impractical or insufficient.

D. EXTENDED COMPARATIVE ANALYSIS

Our DeepRSSI can adequately learn the subtleties of indoor RSSI data to produce highly accurate virtual maps, even at intervals finer than the actual measurement data, or even where the virtual RSSI fingerprint map has not been learned. This approach has different ingenuity from existing augmented or optimization studies on RSSI fingerprint data, summarized in Table 1 and Subsections II-D. In addition to the goal of generating a virtual fingerprint map of DeepRSSI, we experimentally demonstrate a side contribution that also facilitates the augmentation of RSSI data. In this section, we validate our results on other datasets to make comparisons with existing studies in Table 1 using UJIndoorLoc, a publicly available, larger-scale real-world dataset related to state-of-the-art methods, to broaden the scope of experimental comparisons [17]. In two studies by Njima et al., the augmented synthetic fingerprint data show a difference of 3.93 m and 3.47 m in the mean localization error reflecting the average deviation of the predicted location from the actual location of the model [9], [16]. In the study by Ai et al., UJIndoorLoc is used for an average check, and the experiments performed for the augmentation are not compared using the own data set [19]. Our DeepRSSI shows that, when simply augmented for UJIndoorLoc, the mean localization error produces high-quality synthetic data with less error than the original of 2.75m. In addition to generating virtual maps for resource savings, which are the main research objectives, it also demonstrates a very good aspect in

TABLE 4. Comparison of model performance based on the proportion of virtual data with results measured as mean squared error (MSE) in meters [40], [41], [42], [43].

Proportion of virtual data	Random Forest	XGBoost	ElasticNet	KNN	Transformer	ResNet	CNN
0%	32.76	26.12	18.64	32.98	48.36	43.79	49.28
25%	11.16	17.41	16.43	12.60	18.70	8.92	10.20
50%	11.68	8.39	16.40	12.70	19.03	10.06	13.60
75%	11.57	4.94	16.26	12.40	18.91	11.99	12.28
100%	12.91	11.54	16.52	12.50	19.02	8.68	6.01

terms of augmenting similar data targeted by existing studies, demonstrating broader applicability.

V. FUTURE WORK

The industrial significance of our approach lies in its potential to revolutionize RSSI-based mapping generation by enabling conditional generation tailored to specific needs and environments. This capability opens new avenues for enhancing indoor positioning systems, especially in complex and dynamically changing environments, where existing models may struggle to maintain accuracy and reliability. However, despite the promising advances introduced by the DeepRSSI model in the RSSI-based indoor localization domain, this study acknowledges the limitations of no extensive experiments in various real-world environments. Mainly focus on the development and initial verification of the model for generating conditionally generated new maps, and our future research agenda includes extending the experimental verification of the DeepRSSI model to cover a wide range of environments. Future research will aim to conduct a thorough comparative analysis with state-of-the-art indoor localization systems in real-time scenarios. This will not only demonstrate the diversity and robustness of the model, but will also further solidify its industrial relevance by demonstrating its applicability and performance advantages in the operating environment. Through these efforts, we aspire to overcome current limitations and leverage the potential of generative models in enhancing indoor localization techniques for different applications.

VI. CONCLUSION

In this paper, we propose a novel generative model that generates high-quality virtual RSSI fingerprint mapping in an indoor environment and can additionally be used for data augmentation. The comprehensive performance evaluation plan shown in this study provided sufficient evidence for the effectiveness of the proposed model. The indoor localization error analysis revealed that the virtual RSSI data generated by the model can be used effectively in applications such as wireless network optimization and indoor localization. As a result, the work presents an important contribution towards the advancement of the WiFi RSSI fingerprint map. Furthermore, the model for the ability to generate high-quality virtual data has the scalable potential to use data from various communication media, not limited to WiFi RSSI, suggesting a wide range of contributions to significantly saving time and money in generating WiFi fingerprint map and improving performance efficiency. In addition, our model takes advantage of the inherently unpredictable nature of wireless signal propagation to generate diverse and complex RSSI data. Our research proposes a model that can adequately handle data containing such diversity and complexity, and will ultimately improve network performance. In future research, our focus will shift to more complex and extensive real-world scenario testing to evaluate the adaptability and robustness of our model in dynamically changing environments. We intend to improve it to a diverse and reliable generative model through validation of models applying state-of-the-art indoor localization systems and comparative verification through various metrics.

REFERENCES

[1] K. Sung, D. Lee, and H. Kim, "Indoor pedestrian localization using iBeacon and improved Kalman filter," *Sensors*, vol. 18, no. 6, p. 1722, May 2018.

[2] T. Xie, H. Jiang, X. Zhao, and C. Zhang, "A Wi-Fi-based wireless indoor position sensing system with multipath interference mitigation," *Sensors*, vol. 19, no. 18, p. 3983, Sep. 2019.

[3] H. Zhu, L. Cheng, X. Li, and H. Yuan, "Neural-network-based localization method for Wi-Fi fingerprint indoor localization," *Sensors*, vol. 23, no. 15, p. 6992, Aug. 2023.

[4] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 763–776, Jan. 2017.

[5] A. H. Salamah, M. Tamazin, M. A. Sharkas, and M. Khedr, "An enhanced WiFi indoor localization system based on machine learning," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Oct. 2016, pp. 1–8.

[6] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J. Big Data*, vol. 2, no. 1, pp. 1–21, Dec. 2015.

[7] M. Nabati, H. Navidan, R. Shahbazian, S. A. Ghorashi, and D. Windridge, "Using synthetic data to enhance the accuracy of fingerprint-based localization: A deep learning approach," *IEEE Sensors Lett.*, vol. 4, no. 4, pp. 1–4, Apr. 2020.

[8] R. S. Sinha and S.-H. Hwang, "Improved RSSI-based data augmentation technique for fingerprint indoor localisation," *Electronics*, vol. 9, no. 5, p. 851, May 2020.

[9] W. Njima, M. Chafii, A. Chorti, R. M. Shubair, and H. V. Poor, "Indoor localization using data augmentation via selective generative adversarial networks," *IEEE Access*, vol. 9, pp. 98337–98347, 2021.

[10] Z. Tang, S. Li, K. S. Kim, and J. Smith, "Multi-output Gaussian process-based data augmentation for multi-building and multi-floor indoor localization," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2022, pp. 361–366.

[11] R. S. Sinha, S.-M. Lee, M. Rim, and S.-H. Hwang, "Data augmentation schemes for deep learning in an indoor positioning application," *Electronics*, vol. 8, no. 5, p. 554, May 2019.

[12] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," 2017, *arXiv:1712.04621*.

[13] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, Dec. 2022, Art. no. 100258.

[14] A. Poulose and D. S. Han, "Hybrid deep learning model based indoor positioning using Wi-Fi RSSI heat maps for autonomous applications," *Electronics*, vol. 10, no. 1, p. 2, Dec. 2020.

[15] J. H. Seong and D. H. Seo, "Selective unsupervised learning-based Wi-Fi fingerprint system using autoencoder and GAN," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 1898–1909, Mar. 2020.

[16] W. Njima, M. Chafii, and R. M. Shubair, "GAN based data augmentation for indoor localization using labeled and unlabeled data," in *Proc. Int. Balkan Conf. Commun. Netw. (BalkanCom)*, Sep. 2021, pp. 36–39.

[17] J. Torres-Sospedra, R. Montoliu, A. Martínez-Usó, J. P. Avariento, T. J. Arnau, M. Benedito-Bordonau, and J. Huerta, "UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Oct. 2014, pp. 261–270.

[18] L. Serreli, R. Nonnis, G. Bingöl, M. Anedda, M. Fadda, and D. D. Giusto, "Fingerprint-based positioning method over LTE advanced pro signals with GAN training contribute," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, Aug. 2021, pp. 1–5.

[19] H. Ai, T. Hu, and T. Xu, "RAD-GAN: Radio map anomaly detection for fingerprint indoor positioning with GAN," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Nov. 2021, pp. 1–8.

[20] D. J. Suroso, P. Cherntanomwong, and P. Sooraksa, "Deep generative model-based RSSI synthesis for indoor localization," in *Proc. 19th Int. Conf. Electr. Eng./Electron., Comput., Telecommun. Inf. Technol. (ECTI-CON)*, May 2022, pp. 1–5.

[21] D. J. Suroso, P. Cherntanomwong, and P. Sooraksa, "Synthesis of a small fingerprint database through a deep generative model for indoor localisation," *Elektronika Ir Elektrotechnika*, vol. 29, no. 1, pp. 69–75, Feb. 2023.

[22] N. Zhou, X. Zhao, and M. Tan, "Deployment and routing method for fast localization based on RSSI in hierarchical wireless sensor network," in *Proc. IEEE 10th Int. Conf. Mobile Ad-Hoc Sensor Syst.*, Oct. 2013, pp. 614–619.

[23] M. Youssef and A. Agrawala, "The Horus WLAN location determination system," in *Proc. 3rd Int. Conf. Mobile Syst., Appl., Services*, Jun. 2005, pp. 205–218.

[24] K. Kaemarungsi, "Design of indoor positioning systems based on location fingerprinting technique," Ph.D. thesis, School Inf. Sci., Univ. Pittsburgh, Pittsburgh, PA, USA, 2005.

[25] W. Liu, M. Kulin, T. Kazaz, A. Shahid, I. Moerman, and E. De Poorter, "Wireless technology recognition based on RSSI distribution at sub-Nyquist sampling rate for constrained devices," *Sensors*, vol. 17, no. 9, p. 2081, Sep. 2017.

[26] L. Theis, W. Shi, A. Cunningham, and F. Huszár, "Lossy image compression with compressive autoencoders," 2017, *arXiv:1703.00395*.

[27] D. P Kingma and M. Welling, "Auto-encoding variational Bayes," 2013, *arXiv:1312.6114*.

[28] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, Dec. 2019.

[29] A. Jabbar, X. Li, and B. Omar, "A survey on generative adversarial networks: Variants, applications, and training," *ACM Comput. Surveys*, vol. 54, no. 8, pp. 1–49, Nov. 2022.

[30] R. Burks, K. A. Islam, Y. Lu, and J. Li, "Data augmentation with generative models for improved malware detection: A comparative study," in *Proc. IEEE 10th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Oct. 2019, pp. 660–665.

[31] M. Guo, "Attention mechanisms in computer vision: A survey," *Comput. Vis. Media*, vol. 8, no. 3, pp. 331–368, 2022.

[32] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*.

[33] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 1–7.

[34] W.-C. Kang and J. McAuley, "Self-attentive sequential recommendation," in *Proc. IEEE Int. Conf. Data Mining (ICDM)*, Nov. 2018, pp. 197–206.

[35] M. Hahn, "Theoretical limitations of self-attention in neural sequence models," *Trans. Assoc. Comput. Linguistics*, vol. 8, pp. 156–171, Dec. 2020.

[36] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.

[37] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[38] NetSpot. (2023). *Wi-Fi Site Surveys, Analysis, Troubleshooting*. Accessed: Apr. 7, 2023. [Online]. Available: https://www.netspotapp.com/wifi-troubleshooting/wifi-interference.html

[39] M. R. Jivthesh, M. R. Gaushik, P. Adarsh, G. H. Niranga, and N. S. Rao, "A comprehensive survey of WiFi analyzer tools," in *Proc. IEEE 3rd Global Conf. Advancement Technol. (GCAT)*, Oct. 2022, pp. 1–8.

[40] M. Niang, P. Canalda, M. Ndong, F. Spies, I. Dioum, I. Diop, and M. A. A. El Ghany, "An adapted machine learning algorithm based-fingerprints using RLS to improve indoor Wi-Fi localization systems," in *Proc. 4th Int. Conf. Emerg. Trends Electr., Electron. Commun. Eng. (ELECOM)*, Nov. 2022, pp. 1–6.

[41] 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.

[42] R. S. Sinha and S.-H. Hwang, "Comparison of CNN applications for RSSI-based fingerprint indoor localization," *Electronics*, vol. 8, no. 9, p. 989, Sep. 2019.

[43] L. Liu, Q. Zhao, S. Miki, J. Tokunaga, and H. Ebara, "Indoor fingerprinting positioning system using deep learning with data augmentation," *Sensors Mater.*, vol. 34, no. 8, p. 3047, 2022.

[44] D. Nettleton, "Selection of variables and factor derivation," in *Commercial Data Mining*. Boston, MA, USA: Morgan Kaufmann, 2014, pp. 79–104.

[45] Y. Hu, F. Qian, Z. Yin, Z. Li, Z. Ji, Y. Han, Q. Xu, and W. Jiang, "Experience: Practical indoor localization for malls," in *Proc. 28th Annu. Int. Conf. Mobile Comput. Netw.*, Oct. 2022, pp. 82–93.

NAMKYUNG YOON is currently pursuing the Ph.D. degree with the School of Electrical Engineering, Korea University, Seoul, South Korea. He was a Researcher with Korea Electronics Technology Institute (KETI), in 2020. His current research interests include wireless networks, artificial intelligence, and unmanned aerial systems (UAS).

WOOYONG JUNG is currently pursuing the master's degree with the School of Electrical Engineering, Korea University, Seoul, South Korea. His current research interests include wireless networks, artificial intelligence, and unmanned aerial systems (UAS).

HWANGNAM KIM (Member, IEEE) received the B.S.E. degree from Pusan National University, Busan, South Korea, in 1992, the M.S.E. degree from Seoul National University, Seoul, South Korea, in 1994, and the Ph.D. degree in computer science from the University of Illinois at Urbana–Champaign, in 2004. He is currently a Professor with the School of Electrical Engineering, Korea University, Seoul. His research interests include wireless networks, unmanned aerial systems (UAS), UAS traffic management, counter UAS systems, and the Internet of Things.

• • •