UniRM: A Universal Large Model for Multiband 3D Radio Map Construction

Xinyue Jiang, Tong Li, *Member, IEEE*, Zhu Xiao, *Senior Member, IEEE*, Ke Chen, *Member, IEEE*, Shuai Ma, Zhaocheng Wang, *Fellow, IEEE*, and Keqin Li, *Fellow, IEEE*

***Abstract*—Radio maps play a crucial role in optimizing wireless network performance and configuration, providing insights into the spatial distribution of radio frequency signal power. Existing solutions often face challenges in generalizing and adapting across various environments, frequency bands, and vertical dimensions. To overcome these limitations, we propose UniRM, a universal large model designed for constructing multiband 3D radio maps. UniRM leverages large-scale pre-training and prompt learning techniques to accurately generate radio maps across diverse environments, altitudes, and frequency bands. Specifically, UniRM employs a UNet-based encoder-decoder architecture during pre-training to extract universal latent representations that capture shared features across different environmental conditions. A prompt learning module further enhances this by transforming auxiliary inputs, such as environmental descriptions, frequency bands, and altitudes, into discriminative embeddings, thereby enabling effective cross-domain generalization and ensuring robustness in unseen scenarios. Extensive experiments using a large, diverse dataset covering numerous scenarios demonstrate that UniRM outperforms state-of-the-art baselines by over 10% in key metrics, including mean squared error, normalized mean squared error, root mean squared error, and peak signal-to-noise ratio. Notably, zero-shot evaluations highlight UniRM's strong ability to generalize to new environments without retraining. The code for UniRM is available at:https://github.com/Shirleyue/UniRM**

***Index Terms*—Large model, radio map, multiband, prompt learning.**

Received 15 May 2025; revised 10 October 2025; accepted 29 November 2025. Date of publication 4 December 2025; date of current version 20 February 2026. This work was supported in part by the National Natural Science Foundation of China under Grant U22B2057, Grant U24A20247, Grant 62471277, and Grant 62471270; in part by the Young Elite Scientists Sponsorship Program by the China Association for Science and Technology (CAST) under Grant ZB2025-293; in part by Hunan Natural Science Foundation of China under Grant 2025JJ50351; in part by the Key Research and Development Program of Hunan Province under Grant 2024AQ2032; in part by Shenzhen Science and Technology Program under Grant JCYJ20250604185802004; and in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2024A1515030028. *(Corresponding authors: Tong Li; Zhu Xiao; Zhaocheng Wang.)*

Xinyue Jiang, Tong Li, and Zhu Xiao are with the College of Computer Science and Electronic Engineering and the Ministry of Education Key Laboratory of Fusion Computing of Supercomputing and Artificial Intelligence, Hunan University, Changsha 410082, China, and also with Shenzhen Research Institute, Hunan University, Shenzhen 518055, China (e-mail: xyuejiang@hnu.edu.cn; tliay@hnu.edu.cn; zhxiao@hnu.edu.cn).

Ke Chen and Shuai Ma are with the Peng Cheng Laboratory, Shenzhen 518066, China (e-mail: chenk02@pcl.ac.cn; mash01@pcl.ac.cn).

Zhaocheng Wang is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: zcwang@tsinghua.edu.cn).

Keqin Li is with the Department of Computer Science, State University of New York, New Paltz, NY 12561 USA (e-mail: lik@newpaltz.edu).

Digital Object Identifier 10.1109/JSAC.2025.3640150

I. Introduction

THE rapid advancement of wireless networks has introduced a remarkable era of connectivity [1], facilitating a wide range of applications, including the Internet of Things (IoT) [2], vehicular networks [3], low-altitude autonomous aerial vehicle (UAV) networks [4], and evolving cellular systems [5]. To effectively enable these innovative services, thoughtful wireless resource management is essential to optimize spectrum use and promote user experience [6], [7], [8], [9]. Accurate path loss measurements are vital for effective resource management, as they directly influence signal quality and network performance. Unlike traditional methods that rely on costly and time-consuming field surveys, radio map (RM) technology provides a groundbreaking solution by estimating path loss using only location information [10], [11]. This innovation streamlines the utilization of spectrum resources and enhances the adaptability of wireless networks in diverse and dynamic environments. As wireless networks continue to evolve, radio maps are poised to play a pivotal role in unlocking the full potential of next-generation wireless systems, fostering seamless connectivity and enabling more efficient, adaptive, and radio context-aware network management [12], [13], [14].

A radio map is a spatial visualization that captures the distribution of received signal strength or other radio frequency metrics across a given geographical area [15], [16], [17], such as power spectral density, signal-to-noise ratio, or channel characteristics. As shown in Fig. 1, these maps provide a comprehensive understanding of signal propagation behavior, facilitating the design, deployment, and optimization of various wireless networks. Notably, with the rapid proliferation of wireless devices and the increasing demand for heterogeneous services operating across a wide range of frequency bands, there is a growing need for constructing multi-band radio maps [18], [19]. Such maps enable the simultaneous analysis of signal behavior across different frequency bands, which is crucial for enhancing spectrum utilization and facilitating intelligent spectrum management. Moreover, space-air-ground integrated networks have introduced new spatial dimensions to wireless communications, necessitating the development of three-dimensional (3D) radio maps [20], [21], [22]. Unlike traditional two-dimensional representations, 3D radio maps offer a volumetric view of signal propagation, capturing variations in signal characteristics across horizontal and vertical dimensions. This three-dimensional perspective is especially critical for emerging applications such as UAV communica-

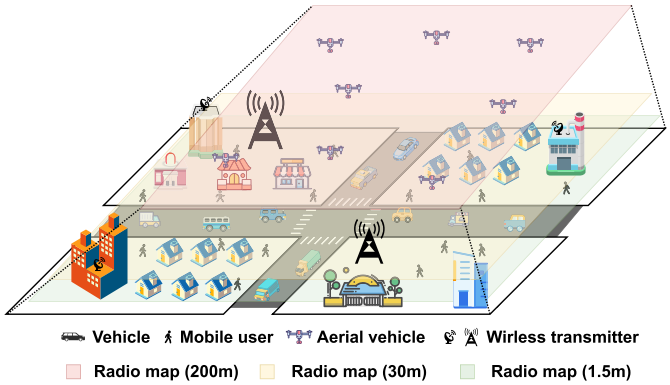


Fig. 1. An illustration of radio maps. 3D radio maps depict the distribution of received signal strength across various heights, offering a comprehensive view of signal propagation behavior. This facilitates the design, deployment, and optimization of cellular, vehicular, and UAV networks.

tions and the strategic deployment of aerial or satellite-based base stations. Therefore, constructing multi-band and 3D radio maps across diverse scenarios has become a key research problem.

With the rapid advancement of deep learning and data-driven techniques, several studies have explored the application of artificial intelligence (AI) models to address the challenge of constructing radio maps. Typically, these methods involve training neural networks that take as input a combination of building layouts, sparse signal strength measurements, and, in some cases, wireless transmitter configurations and output a predicted radio map for a given region. For example, Levie et al. [23] and Jiang et al. [24] proposed UNet-based architectures designed for radio map estimation in urban environments. Similarly, Zhang et al. [25] utilized a generative adversarial network (GAN) to infer radio maps from sparse observations in complex city landscapes. More recently, Wang et al. [26] employed a generative diffusion model, leveraging the strengths of advanced generative modeling for improved accuracy in radio map construction. Despite the promising results achieved by these approaches, existing solutions still face notable limitations in generalizability and adaptability. **First**, their performance is often constrained to specific environments, particularly urban settings. Their limited ability to generalize across diverse scenarios, such as rural, indoor, or mixed terrains, stems from the absence of universal mechanisms for scenario alignment and feature adaptation. **Second**, these models typically focus on a single frequency band and a fixed altitude, overlooking the significant variations in propagation across different frequency bands and vertical dimensions. As a result, *current AI-based approaches fall short of enabling universal 3D radio map construction across diverse environments and the full radio frequency spectrum*. Addressing these limitations is crucial for developing robust, scalable solutions that support future wireless networks, including space-air-ground integrated systems and multi-band, multi-layer communication architectures.

In this paper, we aim to develop a universal model for constructing 3D radio maps across diverse scenarios and frequency bands. In practice, developing such a model poses two fundamental technical challenges. The **first** challenge is designing a general neural network framework capable of handling highly diverse data distributions. Radio propagation patterns vary significantly depending on the environment, such as urban, rural, or indoor settings, as well as the operating frequency band. These variations make it difficult for a small model to generalize effectively across all conditions. The **second** challenge lies in efficiently utilizing and aligning various environmental and metadata information. While auxiliary inputs, such as building layouts, frequency bands, and altitudes, are typically available, they are abstract and non-trivial to integrate in a way that provides meaningful guidance for accurate radio map construction.

To address these challenges, we proposed UniRM, a universal large model for multiband 3D radio map construction, which explores the power of large-scale pre-training and prompt learning to generate accurate radio maps across a wide range of scenarios, altitudes, and frequency bands. We leverage a comprehensive, large-scale dataset encompassing tens of environmental scenarios, frequency bands, and altitudes. Such rich diversity helps the model learn generalized features robust to changes in spatial and spectral domains. Specifically, we adopt a pre-training and prompt learning framework to enhance the scalability and adaptability of UniRM to solve the **first** challenge. A general UNet-based encoder-decoder architecture is employed during pre-training to extract universal latent representations, capturing commonalities across disparate environments. Moreover, we design an innovative prompt learning module that transforms abstract inputs, such as environmental descriptions, frequency bands, and height information, into discriminative prompt embeddings to solve the **second** challenge. This module identifies and aligns underlying shared patterns across domains, dynamically adapting to different distributions and enabling more effective cross-domain generalization. UniRM functions as a universal model capable of reconstructing 3D radio maps across a broad frequency band through these designs. Our approach enables the construction of scalable, accurate, and flexible radio maps in next-generation wireless networks.

Our contributions can be summarized as follows:

- We propose UniRM, a universal large-scale model for constructing multiband 3D radio maps. UniRM harnesses the power of large-scale pre-training and prompt learning to generate accurate radio maps across various environments, altitudes, and frequency bands. To the best of our knowledge, UniRM is the first large model to integrate pre-training and prompt learning within a unified framework designed explicitly for multiband 3D radio map generation.
- UniRM adopts a pre-training and prompt learning framework to enhance scalability and adaptability. During the pre-training phase, a general UNet-based encoder-decoder architecture is employed to learn universal latent representations that capture shared features across various environments. A prompt learning module is designed to incorporate abstract auxiliary information. This module transforms auxiliary information inputs, such as

environmental descriptions, frequency bands, and altitude, into discriminative prompt embeddings. By identifying shared patterns across different domains, the module enables more effective cross-domain generalization, making UniRM robust to diverse and unseen scenarios.

- We conduct extensive experiments to evaluate the generalizability and adaptability of UniRM. To this end, we utilize a comprehensive, large-scale radio map dataset comprising tens of environmental scenarios, frequency bands, and altitudes, including dense urban, ordinary urban, rural, suburban, mountainous, forest, desert, grassland, island, ocean, and lake environments. The experimental results demonstrate that UniRM outperforms state-of-the-art baselines by over 10% in terms of mean squared error (MSE), normalized mean squared error (NMSE), root mean squared error (RMSE), and peak signal-to-noise ratio (PSNR). Furthermore, zero-shot experiments reveal UniRM's remarkable ability to generalize to unseen environments without additional training.

The remainder of this paper is organized as follows. In Section II, we present the related work and preliminaries, in which Section II-C presents the problem formulation. Following that, we elaborate on the methodology in Section III. Section IV provides the evaluation and result discussion. Finally, Section V concludes the paper.

II. Related Work and Preliminaries

In this section, we will provide an overview of existing radio map construction methods, alongside a discussion of large-scale models and prompt learning techniques. Additionally, we present a mathematical formulation of the problem.

A. Radio Map Construction

Traditional radio map construction methods can be broadly classified into physical model-based and interpolation methods. Physical model-based methods rely on predefined signal propagation or channel models [27], [28], [29]. For instance, Lee et al. [30] proposed a log-distance path loss model for Wi-Fi radio map generation. Alternatively, interpolation methods predict unknown values based on the spatial correlation of sparse measurement points without assuming a specific propagation model. A common interpolation technique is the radial basis function (RBF) method [31], which constructs radio maps by fitting observed data using a combination of radially symmetric basis functions, such as Gaussian, multiquadric, or spline functions. Another interpolation approach is inverse distance weighted (IDW) interpolation [32], which relies on spatial distance weights for constructing radio maps. The ordinary Kriging method [33] is also widely used, estimating radio maps by minimizing the variance of the estimation error. Some studies combine physical model assumptions with interpolation techniques to construct radio maps. For example, Bazerque et al. [34] developed an interpolation method grounded in a propagation model to generate radio maps.

The recent success of AI models in wireless networks has sparked a growing interest in AI-based radio map construction [35]. AI-based models are data-driven, automatically learning patterns of signal propagation, which overcome the limitations of fixed-model paradigms in physical model-based approaches. For instance, Li et al. [35] introduced a sparsely self-supervised generative adversarial network (SS-GAN) that generates radio frequency maps from irregularly distributed measurement samples. Similarly, in [23], [24], and [36], UNet-based models were used to construct 2D radio maps in urban environments. Hu et al. [37] proposed Loc-SwinUNet, a deep learning model combining UNet and Swin Transformer architectures, to generate radio maps in urban areas. Liu et al. [38] presented WiFi-Diffusion for achieving fine-grained WiFi radio map estimation using diffusion models. Despite the promising results of AI-based radio map construction methods, they face several critical limitations. First, most models are trained on specific scenarios and struggle to generalize to other scenarios with different propagation characteristics. Second, existing solutions often target isolated frequency ranges, typically sub-6 GHz, neglecting other frequency-dependent propagation physics. Finally, with the advent of 6G wireless networks, 3D coverage is increasingly vital for integrated space-air-ground communications. Motivated by these limitations, we propose leveraging a universal large model to address the complexities of constructing multi-band 3D radio maps.

B. Large Model and Prompt Learning

Large AI models, known as foundation models, have demonstrated exceptional performance across diverse domains, primarily due to their powerful representation learning capabilities and strong generalization capabilities. Foundation models such as GPT [39], BERT [40], and Vision Transformers (ViT) [41] leverage large-scale pre-training to serve as universal feature extractors, enabling effective adaptation to various downstream tasks with minimal fine-tuning. In the field of wireless communications, several large AI models have also been developed to enhance system intelligence and efficiency. For instance, Zhang et al. [42] proposed an integrated air-ground edge-cloud framework that utilizes large models as edge nodes for distributed data processing and small model computation. Similarly, Jiang et al. [43] introduced LAM-SC, a LAM-based semantic communication framework that splits the original image into different semantic segments using universal semantic knowledge. These developments, along with other notable works [44], [45], underscore the growing trend of utilizing large AI models to enhance communication technologies.

Prompt learning has emerged as an effective and efficient approach for leveraging large pre-trained models without requiring full model retraining. This approach has achieved superior performance in large models [46], [47], [48], [49], [50], [51]. Instead of updating the entire model, prompt learning introduces task-specific guidance to steer the universal model toward solving targeted tasks [49], [52], [53], [54]. Our work employs a pre-training and prompting framework to equip the model with general-purpose capabilities, ensuring it is adaptable to diverse scenarios, frequency bands, and heights. This design facilitates transfer learning across various scenarios, enhancing the scalability of multi-band and 3D radio map construction.

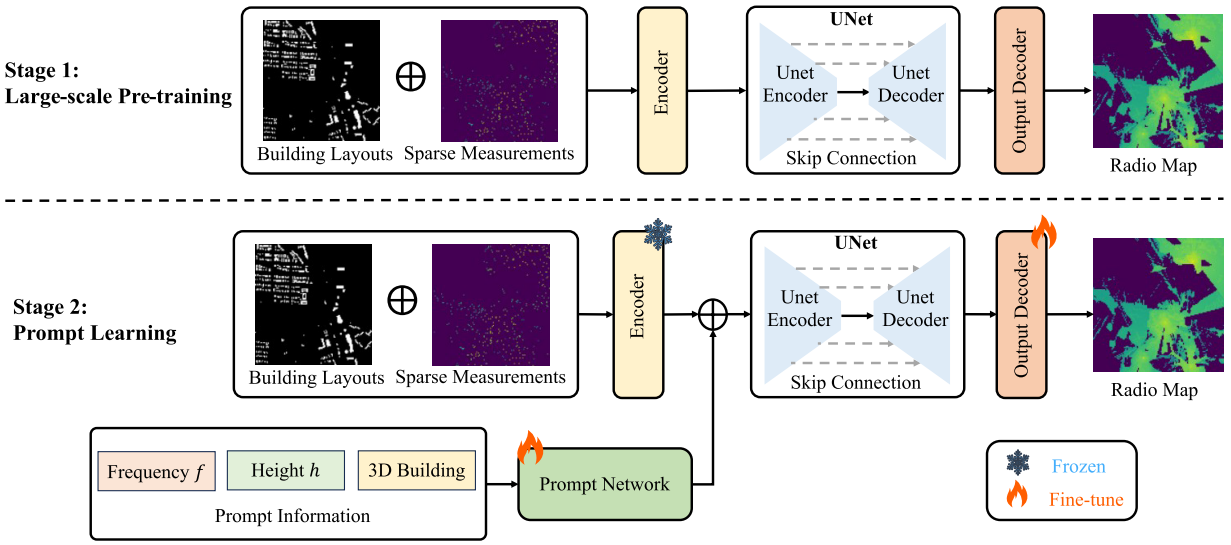


Fig. 2. The overview of the architecture of UniRM, which consists of two stages: (i) Large-scale pre-training. (ii) Prompt learning.

C. Problem Formulation

The radio map construction aims to estimate the fine-resolution radio map from sparse observations. We consider a geographical region uniformly discretized into $W \times L$ grids, with N_t transmitters deployed within the area. Here, $W \times L$ refers to the number of discretized grids. The region is attributed with geographic map with building information, depicted as $\mathbf{B}_{3D} \in \mathbb{R}^{W\times L\times H}$, where $b(x_i, y_i, h_i) = 1, \forall b(x_i, y_i, h_i) \in \mathbf{B}_{3D}$ indicates the existence of a building at grid location (x_i, y_i) and height h_i. Instead of knowing the fine-resolution radio map of the whole area, each region only observes a set of $\mathcal{K} = \{P(x_1, y_1, h_1), P(x_2, y_2, h_2), \cdots, P(x_K, y_K, h_K)\}$ sparse samples collected from sensors. Each sample $P(x_k, y_k, h_k)$ represents the average large-scale receive signal strength within the corresponding grid cell at grid location (x_k, y_k, h_k). This formulation focuses on reconstructing grid-level propagation characteristics from limited measurements.

The objective of multi-band 3D radio map construction is to train a model $\mathcal{D}$ with parameters θ to estimate fine-resolution radio maps at various frequency bands and heights, based on sparse measurements and scenario information. The model $\mathcal{D}$ is trained by minimizing the difference between the reconstructed radio map $\hat{\mathbf{P}}_{f,h} \in \mathbb{R}^{W\times L}$ and the ground truth $\mathbf{P}_{f,h} \in \mathbb{R}^{W\times L}$. The problem of multi-band 3D radio map construction can be formulated as

$$\begin{aligned} \min_{\theta} \quad & \mathcal{L}(\hat{\mathbf{P}}_{f,h}, \mathbf{P}_{f,h}), \\ \text{s.t.} \quad & \hat{\mathbf{P}}_{f,h} = \mathcal{D}_\theta(\mathcal{K}, \mathbf{B}_{3D}, f, h), \end{aligned} \tag{1}$$

where h, f are the height and frequency of interests. $\hat{\mathbf{P}}_{f,h}$ is the estimated radio map at height h with frequency f. $\mathbf{B}_{3D} \in \mathbb{R}^{H\times W\times L}$ is the 3D building information incorporating building projections at different heights.

III. Methodology

In this section, we will provide a detailed explanation of the proposed UniRM. First, we will present an overview of pre-training and the prompt learning framework. After that, we will describe the architectural design of the base model used for large-scale pre-training. Finally, we will explain in the prompt learning stage, including the design and integration with domain-specific knowledge.

A. Pre-Training and Prompt Learning Framework

We aim to propose a universal large model to handle radio map construction across diverse scenarios, frequency bands, and heights, which requires addressing significant deviations of various data distributions. To achieve this, UniRM leverages a pre-training and prompt learning framework, resulting in a universal radio map construction model, UniRM. Fig. 2 illustrates the overview of the UniRM, which consists of two stages as follows.

1) Large-Scale Pre-Training: UniRM is first pre-trained on a large corpus of radio map data encompassing various scenarios, frequency bands, and heights, enabling the model to learn generalizable propagation patterns and transferable feature representations to unseen cases.

2) Prompt Learning: UniRM leverages a prompt network for exploiting the underlying patterns. Prompts are generated based on high-level auxiliary knowledge such as building height, signal frequency, and 3D structural information. These prompts guide the network to adapt its inference process, enabling rapid and efficient adaptation while preserving its generalization capabilities.

B. Base Model

The base model of UniRM is built upon a UNet-style encoder-decoder architecture, enhanced with residual blocks and integrated with a self-attention mechanism to capture both local and global dependencies in radio signal propagation.

1) Encoder-Decoder Structure: As depicted in Fig. 2, the base model consists of three components: (i) a shallow

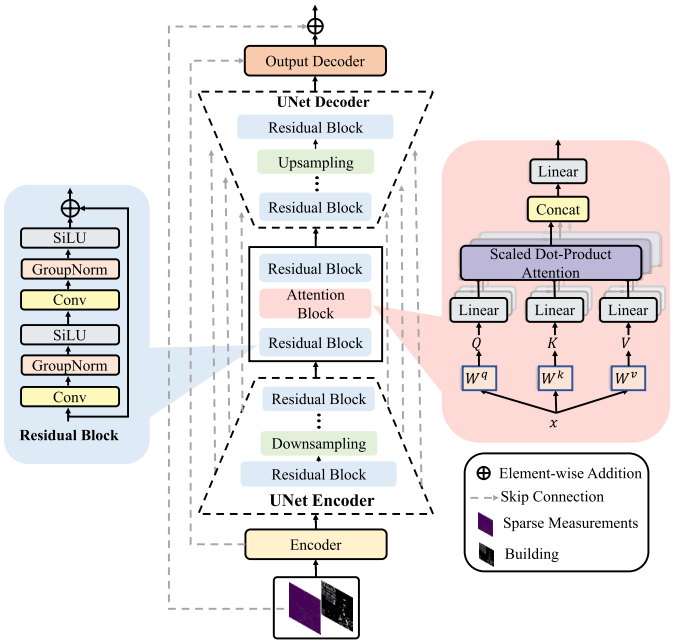


Fig. 3. Detailed structure of the base model in UniRM. The model takes sparse measurements and building layout as input and outputs a fine-resolution radio map. It is built upon a U-Net architecture with residual and attention blocks to enhance spatial feature extraction and long-range dependencies.

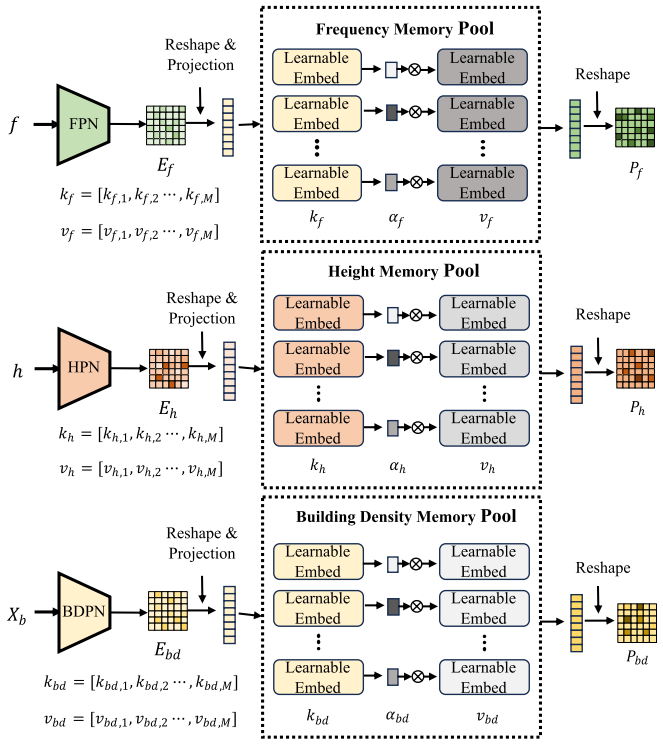


Fig. 4. Illustration of the prompt generation process. Given the frequency f, height h, and building map X_b inputs, UniRM extracts features using three knowledge extraction networks, namely FPN, HPN, and BDPN, and generates prompts via their respective memory pools.

Encoder for input projection, (ii) a deep UNet Encoder-Decoder backbone for hierarchical feature processing, and (iii) a shallow Output Decoder for output construction.

The UNet Encoder-Decoder architecture is depicted in Fig. 3. This symmetric structure captures features at multiple scales while preserving spatial details through skip connections. The encoder progressively reduces spatial resolution to extract hierarchical features, while the decoder reconstructs high-resolution representations through upsampling and the integration of these features. To strengthen representational capacity and support intense networks, we employ residual convolutional blocks [55]. Each residual block contains stacked convolutional layers, followed by Group Normalization [56] and SiLU activation [57], with skip connections to improve gradient flow and training stability.

The Encoder, situated before the UNet backbone, comprises a series of standard convolutional layers that transform raw inputs, i.e., sparse measurements and building layouts, into an embedded feature representation. Similarly, the output decoder, positioned after the UNet, utilizes convolutional layers to transform high-level features into the final dense radio map prediction. This overall design promotes enhanced model expressivity, stable gradient propagation, and improved handling of complex input-to-output mappings.

2) Self-Attention Mechanism: While convolutional layers effectively capture local features, their limited receptive field can restrict the model's ability to understand long-range dependencies, which are crucial for accurately modeling signal propagation in scenarios. To address this, UniRM incorporates a scaled dot-product self-attention block within the encoder-decoder pathway, as illustrated in Fig. 3. The attention mechanism allows the model to compute feature-level dependencies between all spatial locations in the feature map, enhancing its ability to model global interactions:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \tag{2}$$

where Q, K, V are the query, key and value, which can be obtained through the linear transformation of the input features X as follows,

$$Q = XW_Q, \tag{3}$$
$$k = XW_K, \tag{4}$$
$$V = XW_V, \tag{5}$$

where W_Q, W_K, W_V are the projection weights [58]. The self-attention mechanism enables context-aware feature aggregation, allowing spatial locations to attend to semantically relevant regions adaptively. Combining UNet's spatial hierarchy and the attention module's global reasoning ensures effective representation learning for large-scale radio map modeling.

The base model is pre-trained on a large-scale radio map dataset comprising various scenarios, frequency bands, and heights. The training objective minimizes the mean squared error (MSE) between the prediction $\hat{\mathbf{P}}$ and ground truth $\mathbf{P}$:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N}\sum_{i=1}^{N}|\hat{\mathbf{P}}_i - \mathbf{P}_i|^2. \tag{6}$$

C. Prompt Learning

The radio map distributions differ across scenarios, frequency bands, and heights in radio map construction tasks. The key is identifying and aligning underlying related patterns to achieve generalization across data distribution deviation. Prompt learning is a powerful strategy to enhance the generalization of large models by injecting auxiliary knowledge or task-specific cues into the learning process. Instead of retraining the model for each new condition, prompt learning introduces a small set of learnable or crafted prompts that steer the model's behavior under varying inputs or conditions. We adapt this paradigm to radio map construction. As illustrated in Fig. 4, the prompt network takes auxiliary information, i.e., frequency, receiver height, and 3D building density as inputs, and generates distribution-aware prompt embeddings through a memory-based mechanism. These embeddings are fused with encoder features to provide context-aware guidance for radio map construction tasks. This design allows UniRM to adapt to various unseen frequencies, heights, and scenarios.

1) Memory-Based Prompt Generation: The core objective of our prompt generation module is to produce adaptive prompts that enable robust generalization across diverse scenarios, frequency bands, and heights. Unlike using static and pre-defined embeddings, our approach dynamically generates customized prompts tailored to specific input patterns. This adaptability is achieved through memory networks [59] that learn and retrieve domain-specific knowledge.

We maintain a memory pool organized in a key-value structure, where keys K represent prototypical input patterns and values V store the corresponding knowledge representations learned during training. Given an input x, we first project it into the memory space and compute its similarity with all memory keys. The prompt embedding is then generated as an attention-weighted combination of the matched memory values. The memory network can be formulated as follows,

$$\text{Memory}(x) = \sigma(\tanh(W_{proj}x) \cdot K^T) \cdot V, \tag{7}$$

where W_{proj} is a learnable projection matrix, σ is the softmax operator for attention-based weighting, and (K, V) are jointly optimized memory parameters. This process retrieves the most relevant knowledge from the memory pool and fuses it with the original prompt embedding, resulting in a customized, distribution-aware prompt. Thus, inputs with similar propagation patterns are encouraged to generate similar prompts, while distinct inputs obtain specialized prompts tailored to their characteristics. Compared with static prompts, our memory-based design enables UniRM to dynamically adapt to new scenarios, frequency bands, and receiver heights, thereby improving generalization across various radio environments.

2) 3D Building Environment Extraction: Buildings have a profound influence on radio propagation by inducing reflection, diffraction, and shadowing. We define 3D building density as a primary factor in characterizing environment-specific propagation complexity. High-density urban areas typically exhibit rich multipath effects and higher signal attenuation, whereas open environments, such as lakes or plains, display simpler propagation behavior.

We employ a build density prompt network (BDPN) with multi-scale convolutions to capture building density. Convolutions with different kernel scales enable the perception of building information of different sizes and dimensions. The input 3D buildings $X_b \in \mathbb{R}^{C\times W\times H}$ of BDPN are building projections at various heights; each channel of X_b represents the spatial building information at a specific height. The 3D building density extraction process is modeled in Eq.(8).

$$E_{bd} = \sum_{i=1}^{3} \text{Conv}_{k=2^i+1}(X_b), \tag{8}$$

where $\text{Conv}_{k=2^i+1}$ denotes a convolutional layer with a kernel size of $2^i + 1$, and $\sum$ represents feature addition. E_{bd} is the building density extracted by BDPN. Using multi-scale kernels enables BDPN to capture fine-grained and coarse-scale structural variations effectively. The extracted 3D building density reflects the environment-specific complexity of propagation.

The extracted 3D building density, E_bd, is further refined via a memory pool to retrieve and assemble effective prompt embeddings from historical knowledge. The memory pool of building density is defined as follows,

$$M_{bd} = \{(k_{bd,0}, v_{bd,0}), \ldots, (k_{bd,M-1}, v_{bd,M-1})\}, \tag{9}$$

where M is the number of memory slots, all k and v are learnable parameters. Subsequently, useful building prompts are generated based on these optimized memories, which involves using the representations of building density properties as queries to extract valuable memory knowledge, i.e., pertinent embeddings from the memory pool. The overall procedure is illustrated in Fig. 4 and can be formulated as follows,

$$P_{bd} = \sum_{i=1}^{M} \alpha_{bd,i} \cdot v_{bd,i},\ \alpha_{bd,i} = \sigma(E_{bd} \cdot k_{bd,i}^T), \tag{10}$$

where σ is the softmax function for adaptive weights assignment. E_{bd} represents the extracted building knowledge related to the scenario. P_{bd} is the final building prompt embedding.

3) Frequency Propagation Knowledge Extraction: Radio frequency directly affects signal propagation loss, diffraction, and penetration. We transform scalar frequency inputs into physically meaningful representations, utilizing domain knowledge to generalize across different frequency bands. We design a frequency prompt network (FPN) to capture frequency-related latent patterns. FPN consists of two stages. *i)* A scalar frequency is converted into a frequency-space signal loss matrix with physical significance. *ii)* A convolutional encoder to extract spatial-frequency patterns from the estimated propagation loss matrix.

Radio propagation follows path models, such as the Friis transmission equation [60]. The free space path loss can be formulated as,

$$FSPL(f, d) = 20\log_{10}(d) + 20\log_{10}(f) + C, \tag{11}$$

where f is frequency and d is the propagation distance, and C is the unit convert constant. According to Eq.(11), radio propagation loss varies at different frequencies. Due to propagation loss, the received signal strengths vary at different frequencies in the same scenario and at the same location.

Inspired by the propagation law, we combine single-frequency values with the propagation model to generate valuable frequency-based spatial loss prompt information for constructing radio maps. In this way, the abstract frequency information is converted into information of actual physical value. Although the transmitter's location is unknown, signal propagation at different frequencies adheres to the inherent laws of propagation. A rough estimation of the propagation characteristics suffices for model generalization. Thus, we propose a method for calculating the frequency-space propagation loss based on existing sparse samples.

Let $\mathbf{P} \in \mathbb{R}^{W \times L}$ denote the received signal strength matrix in the space, where $P_{i,j}$ denotes the received signal strength at position (x_i, y_j). Given a frequency f and a sparse sample set $\mathcal{K} = \{(x_k, y_k) \mid P(x_k, y_k)\}$, we construct a virtual transmitter set by selecting the top-N maximum power observations:

$$\mathcal{T} = \left\{ \left(x_n^t, y_n^t\right) \mid P_n^{\max} = \arg\max_{\mathcal{K}}^{(n)} P(x, y) \right\}. \tag{12}$$

For each virtual source (x_n^t, y_n^t), we compute relative distance matrix

$$\mathbf{D}_n = \left[\sqrt{(x_i - x_n^t)^2 + (y_i - y_n^t)^2} \right]_{W \times L}. \tag{13}$$

Next, we model frequency-spatial path loss

$$\mathcal{L}(f, \mathbf{D}_n) = 20 \log_{10}(\mathbf{D}_n) + 20 \log_{10}(f) + C, \tag{14}$$

where C is the unit conversion constant. Thus, we integrate the frequency propagation characteristics of N virtual transmitters and generate the frequency-dependent signal strength feature.

$$\tilde{\mathbf{P}} = \frac{1}{N} \sum_{n=1}^{N} (P_n - \mathcal{L}(f, \mathbf{D}_n)), \tag{15}$$

where P_n is the observed received signal strength associated with the n-th virtual transmitter, and $\tilde{\mathbf{P}}$ is the frequency-dependent signal strength feature. Importantly, Eq.(15) is not intended as a strict propagation equation, but rather as a frequency-aware feature transformation that encodes relative attenuation patterns. The feature matrix $\tilde{\mathbf{P}}$ is used as prompt knowledge to enhance UniRM's frequency adaptability.

The process mentioned above converts a given frequency value into a frequency-based spatial signal strength matrix. After that, we adopt a convolution network to capture frequency-spatial features and encode them as frequency-based knowledge, depicted as E_f, to guide radio map construction. The frequency-based knowledge E_f is refined through a memory pool to retrieve and assemble effective prompt embeddings from historical knowledge. The memory pool of frequency propagation knowledge can be defined as follows,

$$M_f = \{(k_{f,0}, v_{f,0}), (k_{f,1}, v_{f,1}), \ldots, (k_{f,M-1}, v_{f,M-1})\}, \tag{16}$$

where M is the number of memory slots, all k and v are learnable parameters. Subsequently, useful frequency prompts are generated based on these optimized memories. This involves using the representations of frequency propagation properties as queries to extract valuable memory knowledge, i.e., pertinent embeddings from the memory pool. Fig. 4 illustrates the process and is formulated as follows,

$$P_f = \sum_{i=1}^{M} \alpha_{f,i} \cdot v_{f,i},\ \alpha_{f,i} = \sigma(E_f \cdot k_{f,i}^T), \tag{17}$$

where E_f represents the extracted knowledge related to frequency. P_f is the final frequency prompt embeddings guiding the construction model.

4) Height Information Extraction: Propagation loss and building density vary with different heights. The given height information is utilized in two aspects. We calculate the propagation loss differences caused by height variations. We define h_0 as the reference height (set to the transmitter altitude). Based on Eq.(11), the propagation loss variation at height h relative to the reference height h_0 can be expressed as

$$\mathcal{L}(h) = 20 \log_{10} \left(\frac{h}{h_0} \right). \tag{18}$$

Thus, when height information is available as prompt knowledge, we augment Eq.(15) to incorporate 3D height-related propagation effects, yielding

$$\tilde{\mathbf{P}} = \frac{1}{N} \sum_{n=1}^{N} (P_n - \mathcal{L}(f, \mathbf{D}_n) - \mathcal{L}(h)). \tag{19}$$

Eq.(19) ensures that variants in propagation loss due to different receiver heights are captured. While it assumes line-of-sight (LOS) conditions, the non-line-of-sight (NLOS) effects caused by building blockage and shadowing are addressed separately through the building density embedding in Eq.(8). Therefore, UniRM integrates both height-dependent and environment-dependent propagation characteristics.

Building information may differ at different heights. We obtain the building information at a specific height and then apply multi-scale convolutions to extract the building density. The height information extraction is formulated as follows:

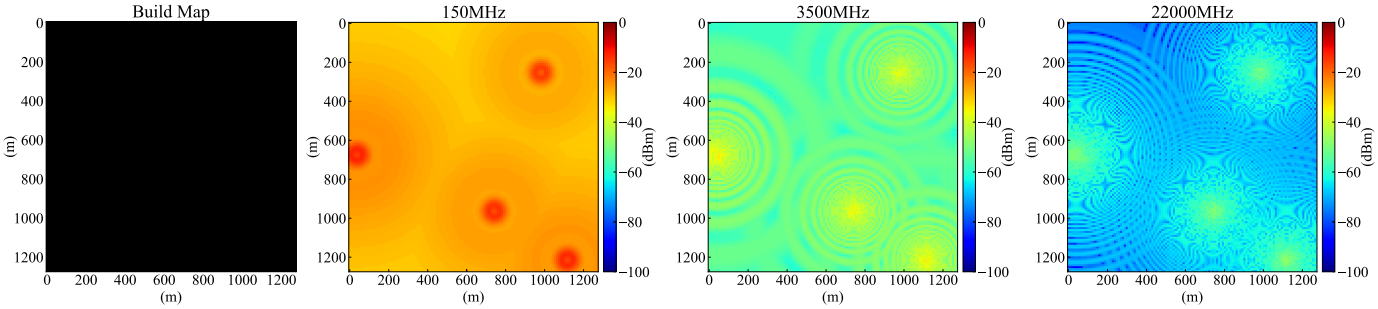
$$E_h = \sum_{i=1}^{3} \text{Conv}_{k=2^i+1}((X_h), \tag{20}$$

where X_h is the building information at height h. This height-aware representation captures both the spatial geometry and the elevation-induced signal variation. The extracted height information is further refined via a memory pool to retrieve and assemble effective prompt embeddings from historical knowledge. The memory pool of height propagation knowledge can be defined as follows:

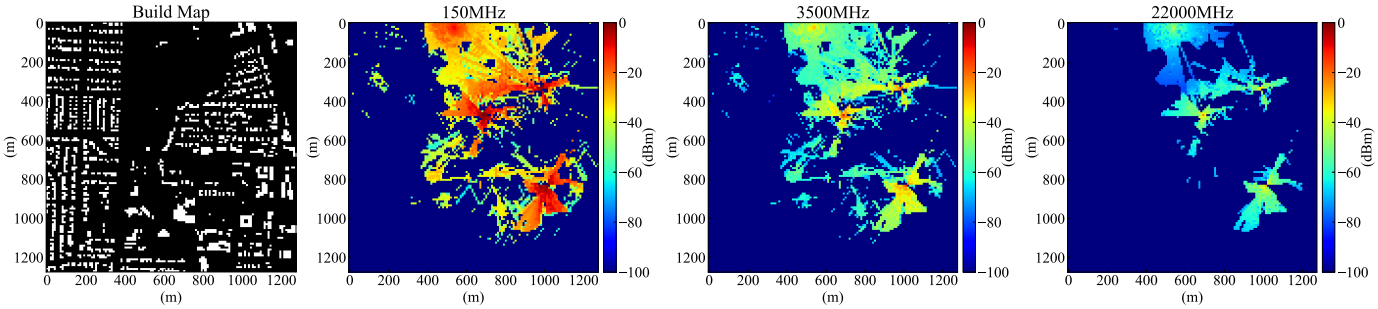
$$M_h = \{(k_{h,0}, v_{h,0}), (k_{h,1}, v_{h,1}), \ldots, (k_{h,M-1}, v_{h,M-1})\}, \tag{21}$$

where M is the number of memory slots, all k and v are learnable parameters. Subsequently, useful height prompts are generated based on these optimized memories. This involves using the representations of height information as queries to extract valuable memory knowledge, i.e., pertinent embeddings from the memory pool. It can be formulated as follows:

$$P_h = \sum_{i=1}^{M} \alpha_{h,i} \cdot v_{h,i},\ \alpha_{h,i} = \sigma(E_h \cdot k_{h,i}^T), \tag{22}$$



(a) A building map and radio maps of different frequency bands in a lake scenario.



(b) A building map and radio maps of different frequency bands in a dense urban scenario.

Fig. 5. Building maps and radio maps of the SpectrumNet dataset across different scenarios and frequencies.

TABLE I
PROPERTIES OF THE SPECTRUMNET DATASET

Properties	SpectrumNet
Real world maps	15300
Projection area of each map	1.28km×1.28km
Spatial resolution	10m
Scenario types	11
Sample heights	1.5m, 30m, 200m
Frequency bands	150MHz, 1.5GHz, 1.7GHz, 3.5GHz, 22GHz

where E_h represents the extracted knowledge related to height. P_h is the final height prompt embeddings guiding the construction model.

The final set of prompts P_{bd}, P_f, P_h is injected into the encoder-decoder architecture (Recall Fig. 2) to provide adaptive guidance for radio map construction. The design of independent prompt pathways allows the model to flexibly recombine scenarios, frequency, and height cues to suit different deployment scenarios and conditions.

IV. PERFORMANCE EVALUATIONS

A. Datasets

We evaluate the performance of the proposed large model UniRM in the SpectrumNet dataset [19]. The SpectrumNet dataset has 94458 radio maps of various scenarios, frequency bands, and heights. As presented in Table I, SpectrumNet comprises 15300 distinct geographical areas, encompassing eleven scenarios: grassland, island, ocean, lake, suburban, dense urban, rural, ordinary urban, desert, mountainous, and forest. The size of each area is 1.28km×1.28km, with a spatial resolution of 10m, resulting in radio map images of size 128×128. The radio maps are provided at three different heights for each area: 1.5m, 30m, and 200m above ground level, reflecting the radio maps for ground nodes, nodes at building heights, and aerial nodes, respectively. Additionally, SpectrumNet encompasses five different bands, ranging from 150 MHz to 22 GHz, which cover most of the working frequencies of wireless networks. In SpectrumNet, the received power values are normalized to the range $[0, 1]$ through linear transformation, where -120 dBm maps to 0 and 60 dBm maps to 1. This normalization facilitates stable training of neural networks. For metric computation and visualization, the values are inverse-normalized back to dBm to ensure that the evaluation is based on physical received signal strength data rather than color representations.

Fig. 5 illustrates examples of two scenarios at different frequencies in the SpectrumNet. As depicted in Fig. 5, different building layouts result in distinct propagation effects in various scenarios. We can observe that the radio map differs between the frequency bands of 150 MHz and 22 GHz.

B. Baselines and Metrics

1) Baselines: We compare UniRM with a broad collection of state-of-the-art models for radio map construction. To ensure a comprehensive comparison, we investigate the proposed UniRM's performance in conjunction with interpolation and deep learning methods.

- **GT** [61]. For radio map construction, we consider the state-of-the-art interpolation-based method, Graph-Interpolation (GT). GT leverages graph signal interpolation techniques to estimate unmeasured radio signal values based on spatial correlations.
- **RadioUNet** [23]. RadioUNet employs a symmetric encoder-decoder architecture that effectively captures

contextual information, making it a representative and widely adopted deep learning-based solution.

- **PMNet** [36]. PMNet is a scalable and generalizable framework for path loss prediction, which adopts an encoder-decoder architecture with atrous (dilated) convolutions to handle scale variations and capture the broader spatial context in radio map data.
- **PEFNet** [24]. PEFNet is a UNet-based method for radio map construction that jointly leverages expert knowledge and data-driven learning.
- **UNetDCN** [62]. UNetDCN is also a framework that employs UNet as its backbone. Based on the standard UNet, UNetDCN replaces some convolutional layers with detachable convolutional layers. This layer calculates the offset of sampling points through learnable parameters to flexibly expand the receptive field, facilitating the propagation of information in any direction.
- **AE** [63]. AE is a radio map construction method based on deep autoencoders, which uses a fully convolutional deep completion autoencoder architecture to exploit the underlying manifold structure of radio maps.
- **RME-GAN** [25]. RME-GAN employs a conditional generative adversarial network (cGAN) to generate radio maps by learning the distribution of radio signals conditioned on known measurements.

2) Metrics: We employ four metrics to evaluate the quality of radio map construction: the mean squared error (MSE), root mean squared error (RMSE), normalized mean squared error (NMSE), and peak signal-to-noise ratio (PSNR).

MSE measures the average of the squares of the errors, i.e., the average squared difference between the estimated values $\hat{y}_i$ and the ground truth y_i. A lower MSE indicates that the predictions are closer to the reference values.

$$\text{MSE} = \frac{1}{N}\sum_{i=0}^{N}(y_i - \hat{y}_i)^2, \tag{23}$$

where N is the number of data points, y_i is the ground truth value, and $\hat{y}_i$ is the predicted value.

RMSE is the square root of MSE, providing an error metric in the same units as the original measurements. It reflects the magnitude of the prediction error and is widely used to evaluate model accuracy.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=0}^{N}(y_i - \hat{y}_i)^2}. \tag{24}$$

NMSE normalizes the squared error by the signal power of the ground truth. It provides a relative error metric that is independent of the absolute scale of the measurements, facilitating fair comparison across different datasets or scenarios.

$$\text{NMSE} = \frac{\sum_{i=0}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=0}^{N} y_i^2}. \tag{25}$$

PSNR is defined as the ratio between the maximum possible power of a signal and the power of interfering noise that affects the fidelity of its representation. PSNR is typically expressed in decibels (dB) and provides an approximate measure of the perceived quality of construction.

$$\text{PSNR} = 10\log_{10}\left(\frac{r^2}{\text{MSE}}\right), \tag{26}$$

where r is the maximal possible signal value in the radio map.

3) Implementation Details: Sparse observations are sampled uniformly at random from each radio map for training and testing. We randomly select 163 sampling points from a radio map of size 128×128, resulting in a sampling rate of 1%. The dataset is divided into training, validation, and testing sets in a 7:1:2 ratio. The training process consists of two phases, both utilizing an AdamW optimizer [64] with a cosine decaying [65] learning rate, starting from 5×10^{-5} and reducing to 1×10^{-6}. In the pre-training stage, we train the backbone network of UniRM without any input of prompt information. In the prompt training stage, we introduce the prompt network and prompt information to guide the model in adapting to various data distributions.

TABLE II

PERFORMANCE COMPARISON OF RADIO MAP CONSTRUCTION. METRICS MARKED WITH ↓ INDICATE BETTER PERFORMANCE WITH LOWER VALUES, WHILE THOSE MARKED WITH ↑ INDICATE BETTER PERFORMANCE WITH HIGHER VALUES. BOLD DENOTES THE BEST RESULTS, AND UNDERLINE DENOTES THE SECOND-BEST RESULTS

Methods	MSE↓	NMSE↓	RMSE↓	PSNR↑
GT	454.280	0.0897	20.85	18.973
RadioUNet	235.872	<u>0.0420</u>	<u>15.10</u>	21.520
PMNet	251.748	0.0450	15.66	21.240
PEFNet	<u>233.604</u>	<u>0.0420</u>	15.12	<u>21.580</u>
UNetDCN	233.928	0.0421	15.12	21.560
RME-GAN	263.088	0.0470	16.02	21.030
AE	296.784	0.0530	17.10	20.500
UniRM	**203.472**	**0.0378**	**13.22**	**23.850**
Improv.	**12.90%**	**10%**	**12.40%**	**10.5%**

C. Overall Performance Evaluation

Table II presents the overall performance comparison of different methods on the dataset. The results show that the GT performs the worst across most evaluation metrics, significantly underperforming compared to all other methods. Our proposed UniRM achieves the best results across all four metrics, i.e., MSE, RMSE, NMSE, and PSNR, with significant improvements over existing methods. Notably, UniRM attains the highest PSNR, indicating that the generated radio maps preserve sharper structural edges and more faithful signal distributions than all baselines. These results validate UniRM as a universal large model with superior accuracy for constructing 3D radio maps.

To further evaluate the quality of radio map construction, we provide a visual comparison among RadioUNet, PEFNet, and our proposed UniRM in a dense urban scenario from the SpectrumNet dataset, as illustrated in Fig. 6. Compared to RadioUNet and PEFNet, UniRM exhibits significant visual improvements, achieving sharper and more precise boundary delineation that accurately reflects the shadowing effects

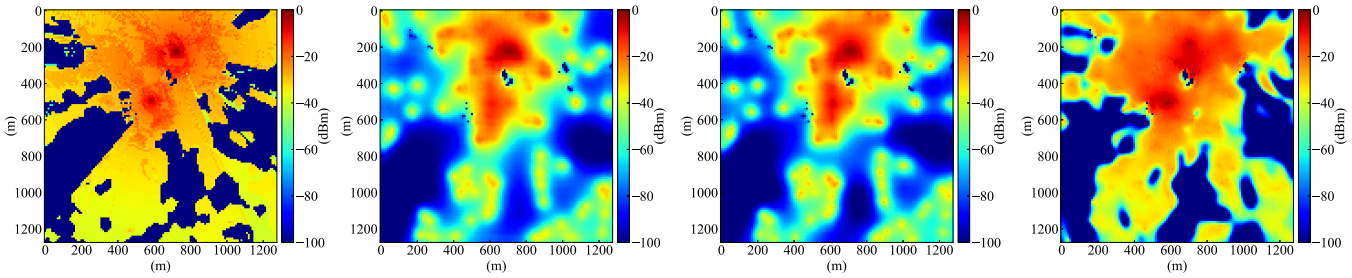


Fig. 6. Visualization: Constructed radio maps of the SpectrumNet dataset. (a) ground truth, (b) radio map constructed by RadioUNet, (c) radio map constructed by PEFNet, (d) radio map constructed by UniRM (Ours).

TABLE III
PERFORMANCE EVALUATION OF RADIO MAP CONSTRUCTION ACROSS VARIOUS SCENARIOS

Methods	Grassland			Island			Ocean			Lake		
	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑
GT	19.188	0.0593	19.598	19.656	0.0672	19.296	5.598	0.0045	30.249	11.430	0.0241	24.747
RadioUNet	14.778	0.0346	22.02	14.400	0.0367	22.04	4.320	0.0027	32.3	9.108	0.0135	26.17
PMNet	14.868	0.0366	21.893	14.832	0.0374	21.73	5.832	0.0049	29.81	9.738	0.0149	25.506
PEFNet	14.562	0.0336	21.921	14.544	0.037	21.961	4.140	0.0024	32.79	9.000	0.013	26.262
UNetDCN	14.598	0.0336	21.89	14.634	0.0369	21.911	4.212	0.0025	32.629	9.072	0.013	26.144
RME-GAN	15.480	0.0376	21.337	15.444	0.0405	21.371	6.336	0.0057	29.097	10.260	0.0165	25.015
AE	16.128	0.0416	21.021	15.894	0.043	21.126	7.758	0.0085	27.327	11.196	0.0222	24.222
UniRM	**13.680**	**0.0337**	**22.882**	**12.600**	**0.0332**	**24**	**3.420**	**0.0019**	**34.651**	**6.997**	**0.0122**	**30.199**

Methods	Dense urban			Rural			Ordinary urban			Desert		
	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑
GT	24.318	0.135	17.413	23.436	0.0899	17.77	24.21	0.115	17.485	24.354	0.1091	17.42
RadioUNet	17.046	0.0657	20.54	17.064	0.0486	20.79	17.892	0.0525	20.137	20.376	0.06	18.96
PMNet	17.712	0.0711	20.177	17.73	0.0517	20.181	17.226	0.0567	20.411	18.162	0.0621	19.995
PEFNet	16.884	**0.0643**	20.596	17.01	0.0487	20.555	16.416	0.0515	20.826	17.838	0.0598	20.177
UNetDCN	16.956	0.0655	20.565	17.208	0.0495	20.448	16.47	0.0515	20.803	17.766	0.0594	20.2
RME-GAN	18.018	0.0732	20.015	16.74	**0.0452**	20.72	18.018	0.0732	20.015	18.018	0.0542	20.061
AE	19.35	0.0842	19.408	19.044	0.0598	19.547	19.206	0.0707	19.472	19.296	0.0693	19.45
UniRM	**16.38**	0.0678	**21.021**	**16.38**	0.0465	**21.02**	**15.75**	**0.0504**	**21.437**	**15.354**	**0.0471**	**21.73**

TABLE IV
PERFORMANCE EVALUATION OF RADIO MAP CONSTRUCTION IN DIFFERENT FREQUENCY BANDS

Methods	150MHz			1.7GHz			3.5GHz			22GHz		
	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑
GT	26.035	0.0746	16.846	20.877	0.0757	18.761	19.500	0.0773	19.355	17.461	0.1325	20.291
RadioUNet	18.828	0.0392	19.68	15.138	0.0400	21.580	14.040	0.0410	22.26	12.042	0.0626	23.54
PMNet	19.494	0.0421	19.374	15.606	0.0425	21.293	14.742	0.0442	21.776	12.384	0.0658	23.277
PEFNet	18.810	0.0393	19.691	15.048	0.0400	21.625	14.058	0.0406	22.226	11.952	0.0617	23.724
UNetDCN	18.792	0.0393	19.706	15.066	0.0399	21.62	14.058	0.0406	22.218	11.790	0.0601	23.594
RME-GAN	19.934	0.0440	19.183	16.012	0.0451	21.085	15.005	0.0458	21.622	12.662	0.0689	23.082
AE	21.174	0.0494	18.642	16.995	0.0504	20.555	15.899	0.0515	21.122	13.500	0.0780	22.517
UniRM	**17.269**	**0.0361**	**20.815**	**13.869**	**0.0367**	**22.632**	**12.944**	**0.0375**	**23.206**	**10.674**	**0.0545**	**24.707**

of urban structures. The estimated radio map generated by UniRM aligns more closely with the ground truth, successfully capturing subtle spatial variations in path loss intensity. These visual results highlight UniRM's superior capability to extract fine-grained radio propagation patterns under complex environmental conditions, which is consistent with the quantitative improvements reported in Table II.

D. Radio Map Construction Across Various Scenarios

We evaluate the performance of radio map construction across various scenarios. For baselines, we train dedicated models for each scenario, since they are not inherently designed to generalize across multiple environments, while UniRM is evaluated across all scenarios. As shown in Table III, we compare the proposed UniRM with baselines in eight scenarios. The performance varies significantly across various scenarios. Radio maps of ocean scenarios achieve the best performance among the ten scenarios. Scenarios with complex buildings, such as dense urban and suburban, yield lower performance. Among all the methods, UniRM outperforms the others in most cases, highlighting its superiority in universal capabilities across scenarios.

E. Multi-Band Radio Map Construction

We evaluate radio map construction across four distinct frequency bands. Each baseline is trained separately for every

TABLE V
PERFORMANCE EVALUATION OF RADIO MAP CONSTRUCTION IN DIFFERENT HEIGHTS

Methods	1.5m			30m			200m		
	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑
GT	31.717	0.2627	15.151	15.854	0.0438	21.476	7.653	0.0099	27.778
RadioUNet	23.040	0.1307	17.900	11.916	0.0226	23.666	5.742	0.0053	30.080
PMNet	23.418	0.1362	17.734	12.366	0.0244	23.322	6.696	0.0070	28.678
PEFNet	22.896	**0.1295**	17.933	11.826	0.0224	23.716	5.562	0.0050	30.367
UNetDCN	22.974	0.1310	17.907	11.698	0.0219	23.812	5.591	0.0050	30.307
RME-GAN	23.775	0.1395	17.605	12.961	0.0268	22.899	7.265	0.0082	27.938
AE	24.944	0.1545	17.184	13.797	0.0303	22.362	8.486	0.0112	26.568
UniRM	**20.844**	0.1318	**19.352**	**9.364**	**0.0217**	**27.094**	**4.623**	**0.0038**	**32.266**

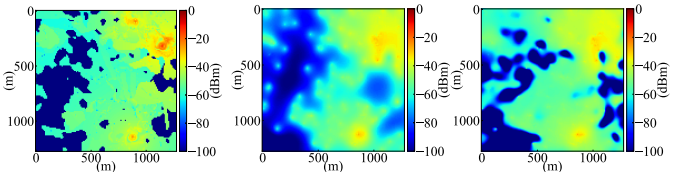


Fig. 7. Visual comparison of generated radio maps using different UniRM variants. (a) ground truth, (b) radio map constructed by UniRM w/o prompt, (c) radio map constructed by UniRM w/prompt.

TABLE VI
ABLATION STUDY ABOUT PROMPT NETWORK. WE USE 'BD', 'F', AND 'H' TO DENOTE BUILDING DENSITY, FREQUENCY, AND HEIGHT PROMPTS, RESPECTIVELY

Methods	MSE↓	NMSE↓	RMSE↓	PSNR↑
UniRM w/o Prompt	218.376	0.0410	14.44	22.169
UniRM w/o 'bd'	204.120	0.0393	13.28	23.790
UniRM w/o 'f'	206.028	0.0395	13.32	23.794
UniRM w/o 'h'	205.092	0.0394	13.27	23.779
UniRM	**203.472**	**0.0378**	**13.22**	**23.850**

frequency band, ensuring a fair comparison with UniRM. Table IV presents the performance results for four distinct frequency bands. The results vary across metrics. While NMSE tends to be lower at lower frequencies, both RMSE and PSNR indicate that construction accuracy is actually better at higher frequencies (e.g., 22 GHz). This discrepancy arises from the physical nature of propagation: low-frequency signals with longer wavelengths undergo stronger diffraction and multipath, which complicates the environment and reduces construction accuracy. In contrast, high-frequency signals exhibit more line-of-sight characteristics, leading to clearer propagation patterns. Despite these variations, UniRM consistently performs best in all frequency bands.

F. Multi-Height Radio Map Construction

We evaluate radio map construction at different receiver heights. For a fair comparison, baselines are trained individually for each height, while UniRM is evaluated across all heights. Table V shows the performance of radio map construction at three distinct heights. The aerial network at 200m achieves the best performance, while the terrestrial network at 1.5m performs the worst. This disparity is attributed to the greater reflection, diffraction, and shadowing effects from buildings on the terrestrial network, making it more challenging to construct a radio map than the aerial network. Our proposed UniRM achieves the best performance across all heights.

G. Ablation Study

Prompts serve as a crucial component in the UniRM model. We next investigate how the designed prompt network contributes to the overall performance. To specify, we use 'bd' to denote building density, 'f' for frequency, and 'h' for height. Our analysis compares the complete design, featuring the prompt network with all three properties, against four degraded versions that individually remove 'bd', 'f', and 'h', and a version without the entire prompt network. Table VI shows the overall results; as one can observe, removing any property decreases performance. By introducing the prompt network, our proposed UniRM model exhibits a superior performance in radio map construction. Fig. 7 presents the qualitative results to illustrate the visual disparity between radio maps generated with and without the prompt network. The qualitative results demonstrate that the prompt network enhances the model's ability to recover edges of the radio map that are obscured by environmental obstacles or propagation attenuation.

H. Generalization Ability Analysis

We evaluate the zero-shot generalization capability of UniRM across diverse scenarios, frequency bands, and heights. We compare our method with three competitive baselines, RadioUNet, PEFNet, and RME-GAN, for comparison.

1) Cross-Scenario Generalization: To assess cross-scenario generalization, we train UniRM on six scenarios from SpectrumNet and evaluate them on four unseen environments (island, lake, ordinary urban, and mountainous). Table VII presents the generalization performance across various scenarios. The proposed UniRM achieves state-of-the-art performance across all unseen test scenarios, demonstrating robust cross-environment generalization.

2) Multi-Band Generalization: For cross-frequency generalization, we train UniRM on three low-frequency bands and test it on two unseen higher-frequency bands, i.e., 3.5 GHz and 22 GHz. Table VIII shows that UniRM maintains robust

TABLE VII
ZERO-SHOT GENERALIZATION PERFORMANCE ACROSS DIFFERENT SCENARIOS

Methods	Island			Lake			Ordinary urban			Mountainous		
	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑
RadioUNet	16.092	0.0441	21.037	10.836	0.0185	24.521	19.445	0.0726	19.370	21.416	0.0905	18.502
PEFNet	16.568	0.0470	20.750	11.210	0.0199	24.253	19.386	0.0717	19.393	21.907	0.0959	18.363
RME-GAN	17.485	0.0525	20.323	11.962	0.0222	23.633	20.989	0.0138	18.716	23.351	0.1065	17.762
UniRM	**13.377**	**0.0366**	**23.550**	**7.398**	**0.0133**	**29.627**	**17.755**	**0.0682**	**20.391**	**19.287**	**0.0796**	**19.649**

TABLE VIII
ZERO-SHOT GENERALIZATION PERFORMANCE ACROSS DIFFERENT FREQUENCY BANDS

Methods	3.5GHz			33GHz		
	RMSE↓	NMSE↓	PSNR↑	RMSE↓	NMSE↓	PSNR↑
RadioUNet	15.232	0.0474	21.504	14.650	0.0920	21.805
PEFNet	15.719	0.0504	21.227	15.037	0.0964	21.575
RME-GAN	15.142	0.0470	21.562	13.695	0.0804	22.397
UniRM	**13.794**	**0.0432**	**22.721**	**11.841**	**0.0679**	**23.842**

TABLE IX
ZERO-SHOT GENERALIZATION PERFORMANCE AT HEIGHT 200 M

Methods	RMSE↓	NMSE↓	PSNR↑
RadioUNet	13.715	0.0226	23.647
PEFNet	13.805	0.0303	22.353
RME-GAN	19.111	0.0574	19.497
UniRM	**9.963**	**0.0208**	**26.577**

performance even on unseen frequency bands, outperforming existing methods.

3) Height Generalization: To evaluate the generalization of UniRM across different heights, we train UniRM at low heights, specifically 1.5m and 30m, and assess its generalization performance at 200m. As depicted in Table IX, UniRM significantly surpasses baselines in high-altitude radio map construction, confirming its strong generalization across varying heights.

The abovementioned experiments validate our knowledge network's ability to extract highly generalizable features and the memory network's capacity to generate optimal prompts for varying conditions. The reason behind this is that by integrating a pre-training and prompt learning framework, UniRM can address the issue of data distribution deviation in various scenarios, frequency bands, and heights.

I. Robustness Analysis

To assess the practical applicability of UniRM, we conduct robustness analyses under various realistic conditions, including varying sampling ratios, noisy environments, and different region sizes. These analyses provide insights into the model's scalability and reliability in practical deployment scenarios.

1) Performance at Different Sampling Ratios: Fig. 8 shows the effect of sampling ratio on performance. As expected, all methods benefit from more measurements, with RMSE decreasing as the sampling ratio increases. Notably, UniRM consistently outperforms all baselines across the entire range. Even at the lowest ratio of 1%, UniRM achieves significantly

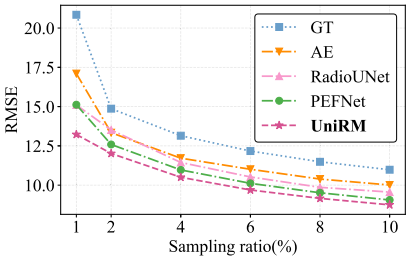


Fig. 8. Performance analysis over different sampling ratios.

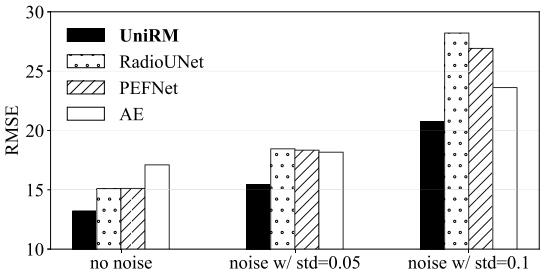


Fig. 9. Performance analysis under noisy conditions. 'std' denotes standard deviation.

TABLE X
PERFORMANCE ON REGIONS OF VARYING SIZES WITH UNIRM TRAINED ON 1.28 KM×1.28 KM RADIO MAPS

Region Sizes	MSE↓	NMSE↓	RMSE↓	PSNR↑
1.28 km×1.28 km	203.472	0.0378	13.22	23.850
1km×1km	251.100	0.0696	15.59	21.400

better accuracy, demonstrating its ability to exploit sparse observations effectively. The performance advantage persists at higher sampling densities, confirming the model's scalability and robustness under varying measurement conditions.

2) Performance Under Noise Conditions: To evaluate the practical robustness of radio map construction models in real-world scenarios, where measurement data are often corrupted by noise due to hardware limitations or environmental factors, we add Gaussian noise with standard deviations of 0.05 and 0.1 to the input measurements. Results in Fig. 9 show that performance degrades for all methods as noise increases. However, UniRM maintains the most stable performance, while baselines such as RadioUNet, PEFNet, and AE exhibit sharper declines at higher noise levels. These results highlight UniRM's superior noise robustness and practical reliability in real-world environments.

TABLE XI
PARAMETER COUNT ANALYSIS OF DIFFERENT MODELS. 'M' AND 'B' DENOTE PARAMETER UNITS IN MILLIONS AND BILLIONS, RESPECTIVELY

Methods	Parameter Count
GT	–
RadioUNet	10.9 M
PMNet	33.4 M
PEFNet	13.3 M
UNetDCN	69.8 M
AE	4.3 M
RME-GAN	5.9 M
LDM	0.49B
DiffusionSAT	0.86B
SpectralGPT	0.63B
UniRM	**0.69 B**

3) Performance on Different Region Sizes: UniRM trained on SpectrumNet using 1.28 km $\times$ 1.28 km radio maps. In real-world scenarios, requirements may vary depending on the area size. To assess the robustness of UniRM to varying region sizes, we evaluate the model, trained on 1.28 km $\times$ 1.28 km radio maps, on smaller regions (1 km $\times$ 1 km). We crop the original radio maps to the new size and pad them to match the model's input dimension (128×128). As shown in Table X, although performance decreases on the 1 km $\times$ 1 km regions, UniRM still achieves satisfactory results, demonstrating its robustness and adaptability to varying region sizes.

J. Parameter Count Analysis

We compare UniRM's parameter count with existing radio map baselines as well as representative foundation models from other domains, i.e., LDM [66] for text-to-image generation, DiffusionSAT [67] for satellite imagery, and SpectralGPT [68] for remote sensing applications. We report the parameter counts in Table XI. UniRM contains 0.69 B parameters, far exceeding previous radio map models, and is comparable to these foundation models. A larger model typically indicates stronger representational capacity and greater ability to capture subtle and complex data patterns. This substantial capacity enables UniRM to capture more complex propagation patterns. Moreover, with pre-training and prompt-learning strategies, UniRM can adapt to varying data distributions and tasks, establishing itself as a general foundation model for radio map construction.

V. CONCLUSION

In this paper, we have developed UniRM, a universal large model for constructing multi-band 3D radio maps. UniRM leverages a pre-training and prompt learning framework that enables modeling to be scalable to various scenarios and frequency bands. We propose a novel prompt network to exploit and align underlying properties of various scenarios, bands, and heights. UniRM achieves STOA performance on a large radio map dataset, validating its universal ability to construct an accurate radio map. UniRM not only establishes a new state-of-the-art in radio map construction but also serves as a scalable and extensible foundation for future wireless communication applications.

UniRM still has limitations, particularly in handling small-scale channel characteristics such as Channel State Information (CSI), which involves more complex spatial-temporal structures and propagation characteristics. The current framework primarily addresses static environments, with limited capability to capture dynamic changes caused by time-varying environments. In the future, we will focus on incorporating spatio-temporal modeling techniques to capture both spatial correlations and temporal dynamics, as well as integrating physical knowledge and environmental dynamics into the learning process. Furthermore, we also plan to expand dataset diversity and scaling model capacity to empower UniRM to evolve toward broader wireless communication applications, including CSI prediction, beamforming optimization, and network planning. We envision UniRM as a universal foundational AI model for wireless environments, enabling efficient and intelligent next-generation communication systems.

REFERENCES

[1] T. Li et al., "Carbon emissions of 5G mobile networks in China," *Nature Sustainability*, vol. 6, no. 12, pp. 1620–1631, Aug. 2023.

[2] Y. Qian, D. Wu, W. Bao, and P. Lorenz, "The Internet of Things for smart cities: Technologies and applications," *IEEE Netw.*, vol. 33, no. 2, pp. 4–5, Mar. 2019.

[3] K. Zheng, Q. Zheng, P. Chatzimisios, W. Xiang, and Y. Zhou, "Heterogeneous vehicular networking: A survey on architecture, challenges, and solutions," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2377–2396, 4th Quart., 2015.

[4] B. Galkin, J. Kibilda, and L. A. DaSilva, "Coverage analysis for low-altitude UAV networks in urban environments," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.

[5] W.-H. Ko, U. Ghosh, U. Dinesha, R. Wu, S. Shakkottai, and D. Bharadia, "EdgeRIC: Empowering real-time intelligent optimization and control in NextG cellular networks," in *Proc. USENIX Symp. Networked Syst. Design Implement. (NSDI)*, 2024, pp. 1315–1330.

[6] D. G. S. Pivoto, F. A. P. de Figueiredo, C. Cavdar, G. R. D. L. Tejerina, and L. L. Mendes, "A comprehensive survey of machine learning applied to resource allocation in wireless communications," *IEEE Commun. Surveys Tuts.*, p. 1, 2025.

[7] J. Zhang et al., "Decision transformers for wireless communications: A new paradigm of resource management," *IEEE Wireless Commun.*, vol. 32, no. 2, pp. 180–186, Apr. 2025.

[8] W. Huang et al., "Safe-NORA: Safe reinforcement learning-based mobile network resource allocation for diverse user demands," in *Proc. 32nd ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2023, pp. 885–894.

[9] J. Gong, Y. Liu, T. Li, J. Ding, Z. Wang, and D. Jin, "STTF: A spatiotemporal transformer framework for multi-task mobile network prediction," *IEEE Trans. Mobile Comput.*, vol. 24, no. 5, pp. 4072–4085, May 2025.

[10] D. Romero, T. N. Ha, R. Shrestha, and M. Franceschetti, "Theoretical analysis of the radio map estimation problem," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 13722–13737, Oct. 2024.

[11] D. Romero and S.-J. Kim, "Radio map estimation: A data-driven approach to spectrum cartography," *IEEE Signal Process. Mag.*, vol. 39, no. 6, pp. 53–72, Nov. 2022.

[12] L. Zhou, H. Mao, X. Deng, J. Zhang, H. Zhao, and J. Wei, "Real-time radio map construction and distribution for UAV-assisted mobile edge computing networks," *IEEE Internet Things J.*, vol. 11, no. 12, pp. 21337–21346, Jun. 2024.

[13] B. Li and J. Chen, "Radio map-assisted approach for interference-aware predictive UAV communications," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 16725–16741, Nov. 2024.

[14] X. Fang et al., "Radio map-based spectrum sharing for joint communication and sensing," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 4541–4558, 2024.

[15] Y. Zeng et al., "A tutorial on environment-aware communications via channel knowledge map for 6G," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 3, pp. 1478–1519, 3rd Quart. 2014.

[16] V.-P. Chowdappa, C. Botella, J. J. Samper-Zapater, and R. J. Martinez, "Distributed radio map reconstruction for 5G automotive," *IEEE Intell. Transp. Syst. Mag.*, vol. 10, no. 2, pp. 36–49, Summer. 2018.

[17] W. Feng et al., "Radio map-based cognitive satellite-UAV networks towards 6G on-demand coverage," *IEEE Trans. Cognit. Commun. Netw.*, vol. 10, no. 3, pp. 1075–1089, Jun. 2024.

[18] X. Li et al., "RadioGAT: A joint model-based and data-driven framework for multi-band radiomap reconstruction via graph attention networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 17777–17792, Nov. 2024.

[19] S. Zhang et al., "Generative AI on SpectrumNet: An open benchmark of multiband 3-D radio maps," *IEEE Trans. Cognit. Commun. Netw.*, vol. 11, no. 2, pp. 886–901, Apr. 2025.

[20] T. Hu, Y. Huang, J. Chen, Q. Wu, and Z. Gong, "3D radio map reconstruction based on generative adversarial networks under constrained aircraft trajectories," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 8250–8255, Jun. 2023.

[21] Z. Chen, H. Wang, and D. Guo, "3-D radio map estimation based on active measurement trajectory selection," *IEEE Wireless Commun. Lett.*, vol. 14, no. 7, pp. 1884–1888, Jul. 2025.

[22] L. Zhao, Z. Fei, X. Wang, J. Luo, and Z. Zheng, "3D-RadioDiff: An altitude-conditioned diffusion model for 3D radio map construction," *IEEE Wireless Commun. Lett.*, vol. 14, no. 7, pp. 1969–1973, Jul. 2025.

[23] R. Levie, Ç. Yapar, G. Kutyniok, and G. Caire, "RadioUNet: Fast radio map estimation with convolutional neural networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 4001–4015, Jun. 2021.

[24] F. Jiang, T. Li, X. Lv, H. Rui, and D. Jin, "Physics-informed neural networks for path loss estimation by solving electromagnetic integral equations," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 15380–15393, Oct. 2024.

[25] S. Zhang, A. Wijesinghe, and Z. Ding, "RME-GAN: A learning framework for radio map estimation based on conditional generative adversarial network," *IEEE Internet Things J.*, vol. 10, no. 20, pp. 18016–18027, Oct. 2023.

[26] X. Wang et al., "RadioDiff: An effective generative diffusion model for sampling-free dynamic radio map construction," *IEEE Trans. Cognit. Commun. Netw.*, vol. 11, no. 2, pp. 738–750, Apr. 2025.

[27] E. Rastorgueva-Foi et al., "Millimeter-wave radio SLAM: End-to-end processing methods and experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 9, pp. 2550–2567, Sep. 2024.

[28] S. Bi, J. Lyu, Z. Ding, and R. Zhang, "Engineering radio maps for wireless resource management," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 133–141, Apr. 2019.

[29] R. Amiri et al., "Indoor environment learning via RF-mapping," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1859–1872, Jun. 2023.

[30] M. Lee and D. Han, "Voronoi tessellation based interpolation method for Wi-Fi radio map construction," *IEEE Commun. Lett.*, vol. 16, no. 3, pp. 404–407, Mar. 2012.

[31] J. Krumm and J. Platt, "Minimizing calibration effort for an indoor 802.11 device location measurement system," Microsoft Res., Redmond, WA, USA, Tech. Rep. MSR-TR-2003-82, Jan. 2003.

[32] S.-P. Kuo and Y.-C. Tseng, "Discriminant minimization search for large-scale RF-based localization systems," *IEEE Trans. Mobile Comput.*, vol. 10, no. 2, pp. 291–304, Feb. 2011.

[33] G. Boccolini, G. Hernández-Peñaloza, and B. Beferull-Lozano, "Wireless sensor network for spectrum cartography based on Kriging interpolation," in *Proc. IEEE 23rd Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Sydney, NSW, Australia, Sep. 2012, pp. 1565–1570.

[34] J. A. Bazerque, G. Mateos, and G. B. Giannakis, "Group-lasso on splines for spectrum cartography," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4648–4663, Oct. 2011.

[35] Z. Li, J. Cao, H. Wang, and M. Zhao, "Sparsely self-supervised generative adversarial nets for radio frequency estimation," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2428–2442, Nov. 2019.

[36] J.-H. Lee and A. F. Molisch, "A scalable and generalizable pathloss map prediction," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 17793–17806, Nov. 2024.

[37] M. Hu, S. Xu, Q. Liu, M. Katwe, C. Yuen, and S. G. Razul, "Locswinunet: A neural network for urban wireless localization using TOA and RSS radio maps," in *Proc. IEEE 33rd Int. Workshop Mach. Learn. Signal Process. (MLSP)*, Sep. 2023, pp. 1–6.

[38] Z. Liu, S. Zhang, Q. Liu, H. Zhang, and L. Song, "WiFi-diffusion: Achieving fine-grained WiFi radio map estimation with ultra-low sampling rate by diffusion models," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 11, pp. 3796–3812, Nov. 2025.

[39] A. Radford et al., "Language models are unsupervised multitask learners," *OpenAI Blog*, vol. 1, no. 8, p.9, 2019.

[40] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol.*Minneapolis, MI, USA: Association for Computational Linguistics, vol. 1, Jun. 2019, pp. 4171–4186.

[41] A. Dosovitskiy et al., "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2020, pp. 1–21.

[42] S. Zhang et al., "Large models for aerial edges: An edge-cloud model evolution and communication paradigm," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 1, pp. 21–35, Jan. 2025.

[43] F. Jiang et al., "Large AI model-based semantic communications," *IEEE Wireless Commun.*, vol. 31, no. 3, pp. 68–75, Jun. 2024.

[44] F. Jiang et al., "Large language model enhanced multi-agent systems for 6G communications," *IEEE Wireless Commun.*, vol. 31, no. 6, pp. 48–55, Dec. 2024.

[45] F. Jiang et al., "Large AI model empowered multimodal semantic communications," *IEEE Commun. Mag.*, vol. 63, no. 1, pp. 76–82, Jan. 2025.

[46] M. Jia et al., "Visual prompt tuning," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Cham, Switzerland: Springer, 2022, pp. 709–727.

[47] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig, "Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing," *ACM Comput. Surv.*, vol. 55, no. 9, pp. 1–35, Sep. 2023.

[48] S. Qiao et al., "Reasoning with language model prompting: A survey," in *Proc. 61st Annu. Meeting Assoc. Comput. Linguistics (Long Papers)*, vol. 1, 2023, pp. 5368–5393.

[49] S. Shen et al., "Multitask vision-language prompt tuning," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2024, pp. 5644–5655.

[50] H. Yao, R. Zhang, and C. Xu, "Visual-language prompt tuning with knowledge-guided context optimization," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 6757–6767.

[51] Y. Yao, A. Zhang, Z. Zhang, Z. Liu, T.-S. Chua, and M. Sun, "CPT: Colorful prompt tuning for pre-trained vision-language models," *AI Open*, vol. 5, pp. 30–38, Jan. 2024.

[52] K. Sohn et al., "Visual prompt tuning for generative transfer learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 19840–19851.

[53] Y. Gao et al., "Visual prompt tuning for test-time domain adaptation," 2022, *arXiv:2210.04831*.

[54] J. Zhu, S. Lai, X. Chen, D. Wang, and H. Lu, "Visual prompt multi-modal tracking," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 9516–9526.

[55] 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.

[56] Y. Wu and K. He, "Group normalization," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.

[57] S. Elfwing, E. Uchibe, and K. Doya, "Sigmoid-weighted linear units for neural network function approximation in reinforcement learning," *Neural Netw.*, vol. 107, pp. 3–11, Nov. 2018.

[58] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2025, pp. 5998–6008.

[59] S. Sukhbaatar, A. Szlam, J. Weston, and R. Fergus, "End-to-end memory networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015, pp. 2440–2448.

[60] H. T. Friis, "A note on a simple transmission formula," *Proc. IRE*, vol. 34, no. 5, pp. 254–256, May 1946.

[61] A. E. C. Redondi, "Radio map interpolation using graph signal processing," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 153–156, Jan. 2018.

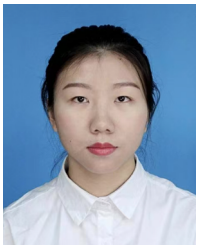
[62] F. Jaensch, G. Caire, and B. Demir, "Radio map estimation—An open dataset with directive transmitter antennas and initial experiments," 2024, *arXiv:2402.00878*.

[63] Y. Teganya and D. Romero, "Deep completion autoencoders for radio map estimation," *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 1710–1724, Mar. 2022.

[64] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2017, pp. 1–18. [Online]. Available: https://api.semanticscholar.org/CorpusID:53592270

[65] I. Loshchilov and F. Hutter, "SGDR: Stochastic gradient descent with warm restarts," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2016, pp. 1–16. [Online]. Available: https://openreview.net/forum?id=Skq89Scxx

[66] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 10674–10685.
[67] S. Khanna et al., "DiffusionSat: A generative foundation model for satellite imagery," in *Proc. ICLR*, 2023, pp. 1–19.
[68] D. Hong et al., "SpectralGPT: Spectral remote sensing foundation model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 8, pp. 5227–5244, Aug. 2024.



Xinyue Jiang received the M.S. degree in communication engineering from Central South University, Changsha, China, in 2021. She is currently pursuing the Ph.D. degree with the College of Computer Science and Electronic Engineering, Hunan University, Changsha, China. Her current research interests include deep learning (DL), jamming recognition, signal processing, and wireless communications.

Tong Li (Member, IEEE) received the Ph.D. degree in computer science and engineering from The Hong Kong University of Science and Technology in 2021 and the Ph.D. degree in computer science from the University of Helsinki in 2022. He is currently a Full Professor with the College of Computer Science and Electronic Engineering, Hunan University, Changsha, China. His research interests include mobile computing, wireless network digital twins, network simulation and optimization, and data-driven networking.

Zhu Xiao (Senior Member, IEEE) received the M.S. and Ph.D. degrees in communication and information systems from Xidian University, China, in 2007 and 2009, respectively. From 2010 to 2012, he was a Research Fellow with the Department of Computer Science and Technology, University of Bedfordshire, U.K. He is currently a Full Professor with the College of Computer Science and Electronic Engineering, Hunan University, China. His research interests include intelligent information processing, the Internet of Vehicles, and wireless communications systems. He is serving as an Associate Editor for IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS.

Ke Chen (Member, IEEE) is currently an Associate Research Fellow with the Peng Cheng Laboratory (PCL). Before joining PCL, he was an Associate Professor with the School of Electronic and Information Engineering, South China University of Technology, China; and a Post-Doctoral Research Fellow with the Department of Signal Processing, Tampere University of Technology, Finland. His research interests include computer vision, pattern recognition, and spectrum sensing.

Shuai Ma received the B.S. and Ph.D. degrees in communication and information systems from Xidian University, Xi'an, China, in 2009 and 2016, respectively. From 2014 to 2015, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA. From 2016 to 2019, he was an Associate Professor with the School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China. From 2019 to 2022, he was a Post-Doctoral Fellow with Telecom Paris, France. Since 2023, he has been an Associate Researcher with the Peng Cheng Laboratory, Shenzhen, China. His research interests include semantic communications, visible light communications, and network information theory.

Zhaocheng Wang (Fellow, IEEE) received the B.S., M.S., and Ph.D. degrees from Tsinghua University, in 1991, 1993, and 1996, respectively. From 1996 to 1997, he was a Post-Doctoral Fellow with Nanyang Technological University, Singapore. From 1997 to 1999, he was a Research Engineer/Senior Engineer with OKI Techno Centre (Singapore) Pte. Ltd., Singapore. From 1999 to 2009, he was a Senior Engineer/Principal Engineer with Sony Deutschland GmbH, Germany. Since 2009, he has been a Professor with the Department of Electronic Engineering, Tsinghua University. His research interests include wireless communications, millimeter wave communications, optical wireless communications, and AI empowered wireless communications.

Keqin Li (Fellow, IEEE) received the B.S. degree in computer science from Tsinghua University in 1985 and the Ph.D. degree in computer science from the University of Houston in 1990. He is a SUNY Distinguished Professor with the State University of New York and the National Distinguished Professor with Hunan University, China. He has authored or co-authored more than 1200 journal articles, book chapters, and refereed conference papers. He holds nearly 80 patents announced or authorized by Chinese National Intellectual Property Administration. He is among the world's top few most influential scientists in parallel and distributed computing, regarding single-year impact (ranked No. 2) and career-long impact (ranked No. 3) based on a composite indicator of the Scopus citation database. He is listed in Scilit Top Cited Scholars (2023–2025) and among the top 0.02% out of over 20 million scholars worldwide based on top-cited publications in the last ten years. He is listed in ScholarGPS Highly Ranked Scholars (2022–2025) and is among the top 0.002% out of over 30 million scholars worldwide based on a composite score of three ranking metrics for research productivity, impact, and quality in the recent five years. He is an AAAS Fellow, an AAIA Fellow, an ACIS Fellow, and an AIIA Fellow. He is a member of European Academy of Sciences and Arts. He is a member of the Academia Europaea (Academician of the Academy of Europe). He received the IEEE TCCLD Research Impact Award from the IEEE CS Technical Committee on Cloud Computing in 2022 and the IEEE TCSVC Research Innovation Award from the IEEE CS Technical Community on Services Computing in 2023. He won the IEEE Region 1 Technological Innovation Award (Academic) in 2023. He was a recipient of the 2022–2023 International Science and Technology Cooperation Award, and the 2023 Xiaoxiang Friendship Award of Hunan Province, China. He is a member of the SUNY Distinguished Academy.