

Article

Reconstruction of Radio Environment Map Based on Multi-Source Domain Adaptive of Graph Neural Network for Regression

Xiaomin Wen ¹, Shengliang Fang ^{2,*} and Youchen Fan ²

¹ Graduate School, School of Space Information, Space Engineering University, Beijing 101416, China; eeiwxm@163.com

² School of Space Information, Space Engineering University, Beijing 101416, China; love193777@sina.com

* Correspondence: eeiwsl@163.com; Tel.: +86-1538-471-2985

Abstract: The graph neural network (GNN) has shown outstanding performance in processing unstructured data. However, the downstream task performance of GNN strongly depends on the accuracy of data graph structural features and, as a type of deep learning (DL) model, the size of the training dataset is equally crucial to its performance. This paper is based on graph neural networks to predict and complete the target radio environment map (REM) through multiple complete REMs and sparse spectrum monitoring data in the target domain. Due to the complexity of radio wave propagation in space, it is difficult to accurately and explicitly construct the spatial graph structure of the spectral data. In response to the two above issues, we propose a multi-source domain adaptive of GNN for regression (GNN-MDAR) model, which includes two key modules: (1) graph structure alignment modules are used to capture and learn graph structure information shared by cross-domain radio propagation and extract reliable graph structure information for downstream reference signal receiving power (RSRP) prediction task; and (2) a spatial distribution matching module is used to reduce the feature distribution mismatch across spatial grids and improve the model's ability to remain domain invariant. Based on the measured REMs dataset, the comparative results of simulation experiments show that the GNN-MDAR outperforms the other four benchmark methods in accuracy when there is less RSRP label data in the target domain.



Citation: Wen, X.; Fang, S.; Fan, Y. Reconstruction of Radio Environment Map Based on Multi-Source Domain Adaptive of Graph Neural Network for Regression. *Sensors* **2024**, *24*, 2523. <https://doi.org/10.3390/s24082523>

Academic Editor: Boon-Chong Seet

Received: 3 March 2024

Revised: 6 April 2024

Accepted: 10 April 2024

Published: 15 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: graph neural networks; invariant inference; latent space; multi-source domain adaptation; radio environment map; spatial distribution matching

1. Introduction

In recent years, with the rapid development of wireless communication technology and mobile internet, wireless spectrum resources have been demanded increasingly, and the available frequency bands have been allocated. Therefore, it is particularly important to improve the efficiency of spectrum dynamic utilization [1]. Cognitive Radio (CR) is considered one of the solutions to solve the contradiction between spectrum supply and demand [2]. Moreover, the reconstruction of the REM is a crucial technology for CR. REM reconstruction [3] refers to the process of establishing the distribution of spectrum data in the spatial domain of the entire cognitive radio network based on the limited spatial spectrum data obtained by spectrum monitoring stations.

The methods for reconstructing REM can be roughly divided into two categories [4]: the radiation source parameter estimation method and the electromagnetic wave data statistical method. The first type of method tends to focus more on the exploration of path loss prediction models in radio wave propagation, which is not our research content in this paper. The second type of method commonly includes spatial interpolation algorithms based on data sparsity (such as Kriging [5], Radial Basic Functions (RBF) [6], and Spline [7]), and tensor completion methods based on low-rank assumptions [8–10]. The above methods

all rely on assumptions such as the smoothness, sparsity, and low rank of REM in space. However, the spatial propagation of electromagnetic waves is a complex process, and those assumptions cannot achieve an accurate prediction of REM. Researchers have considered data-driven deep neural networks (DNNs) [11], which can learn complex underlying structures from data to help address these limitations and challenges. With inspiration from the powerful image generation capability of the Generative Adversarial Network (GAN), the REM was transformed into an image and input into a wireless map estimation model designed based on the GAN [12,13]. However, these methods do not take into account the complex fading effects of wireless transmission environments, such as multipath effects. Although they consider geographic information near CR devices, obstacles or scatterers far away from receiving devices may also significantly affect the receiving power. A model based on convolutional neural networks (CNNs) was applied to wireless map construction in [14,15]. The author drew inspiration from the idea of image processing, and treated obstacles that affect radio wave propagation as a mask matrix added to the network input. This processing method has the disadvantage of uncertainty in shadow fading effects in some scenarios, such as non-line-of-sight (NLOS) propagation.

Given the advantages of GNN in processing unstructured data tasks, research on graph-structured data modeling in the field of wireless communication [16,17] has also made beneficial explorations in recent years. In [18,19], wireless systems are modeled as stochastic wireless graph models that are subject to interference and fading effects. The Random Edge Graph Neural Network (REGNN) is proposed to solve constrained power control problems. Additionally, ref. [20] utilizes GCN technology to design an algorithm for traffic prediction in cellular systems. To solve the difficulty of shadow estimation under NLOS conditions, ref. [21] divides REM into several sub-regions and considers them as nodes of graph-structured data. The received signal strength measurement values are expressed by whether there are edges between nodes to represent radio line-of-sight (LOS)/NLOS propagation, and the label data values of nodes represent spatial distribution. In [22], the author converted each pixel into a node and constructed a graph structure based on the “proximity similarity principle” and “ray tracing model” to illustrate the edge relationship between the receiver and transmitter.

Cognitive radio networks (CRN) can generally be regarded as a complex network topology structure composed of several nodes (such as wireless spectrum sensing devices or base stations) and wireless communication links between them. As previously mentioned, traditional methods have difficulty handling wireless communication networks from the perspective of graph structure modeling. However, graph representation learning methods based on GNN can effectively mine the structural and feature information in graph data [1], to capture the structural and dependency relationships between nodes. This approach enables a more precise modeling of the topology of wireless networks more accurately. Motivated by the above observations, we will carry out REM reconstruction research based on GNN.

However, GNN relies heavily on the input graph structure data, and the accuracy of the relationships between nodes in the graph data is a key factor affecting the robustness and universality of the task model. From the above research, it can be observed that, on the one hand, manually defining the graph structure data unavoidably introduces noise information, resulting in the presence of excessive or missing edges. On the other hand, although REM data contains implicit structural information, the spatial complexity of radio propagation prevents its direct conversion into explicit graph-structured data. These two issues contribute to the susceptibility of noise data and inaccurate graph structure data when applying GNN directly to REM construction, ultimately leading to a decrease in the performance of the GNN model.

Meanwhile, a common challenge faced by DL is also present in the REM prediction task, which is the significant human and resource cost involved in collecting a sufficient amount of valid electromagnetic spectrum-labeled data for training network models. Domain adaptation (DA) [23] aims to transfer models from labeled source domains to unlabeled

beled or sparsely labeled target domains. By leveraging DA techniques, we can effectively reduce the data collection cost associated with REM annotation. However, unlike the effectiveness of DA algorithms in classification tasks, most DA algorithms struggle to effectively address regression problems [24]. This is because the continuity of the output space in regression models results in the absence of clear decision boundaries.

To address the challenges of graph structure noise and insufficient spatial spectrum data for target REM completion, we propose a REM reconstruction method based on Graph Neural Network Multi-Source Domain Adaptation Regression (GNN-MDAR). This approach aims to achieve robustness and regression prediction for cross-domain REM spatial graph structure data. Specifically, the proposed GNN-MDAR model makes the following contributions:

- (1) Introducing multiple different spatial REMs to enhance training data also leads to cross-domain drift issues in graph structure data statistical characteristics. To address this, we introduce the idea of variational graph structure learning into multi-domain adaptation algorithms and design a cross-domain graph structure (GS) alignment module based on the theory of variational information bottleneck. This module is used to learn the spatial graph structure shared information of grid features in source and target REMs.
- (2) In the process of multi-source domain adaptation learning, to avoid the problem of suppressing target domain task performance caused by the forced migration of low-correlation grid features from the source domain, we also designed a spatial distribution matching module. This module achieves alignment of source and target domain grid features in the latent space, capturing the domain invariance of cross-domain REM grids. It enhances the generalization capability of the proposed model for predicting RSRP values in target REM grids.
- (3) We constructed a semi-supervised learning loss function related to the multi-domain adaptive (MDA) REM prediction task. Specifically, we used grid data with RSRP values from source and target REMs to construct the supervised loss function, ensuring the consistency of the trained model with the given label data. We also used grid data without RSRP values from the target REM to construct a semi-supervised loss function to force the regression model to smoothly fit the RSRP prediction data.

The remainder of this paper is organized as follows. Section 2 reviews the relevant research on MDA regression tasks. Section 3 formally defines the problem of wireless environment map prediction based on GNN in this paper. Section 4 introduces the specific principles and methods of the proposed model. Section 5 evaluates the effectiveness of GNN-MDAR through comparative analysis with baseline models in simulation experiments. Finally, Section 6 summarizes and discusses the research work conducted in this paper.

2. Related Works

2.1. Distribution Matching

The idea of distribution matching is a statistical technique used for aligning distributions and can effectively solve the problem of imbalanced data distribution. Two representative techniques of distribution matching, namely Variational Inference (VI) [25] and Adversarial Learning [26], have been extensively studied in various applications such as human pose estimation [27], few-shot learning [28], long-tail recognition [29], and tracking [30]. VI flexibly matches various forms of distributions for the objective function through probabilistic calculation, and can utilize more complex higher-order information. Generative adversarial learning includes a generator and a discriminator, usually requiring a large amount of computing resources and taking a long time to train. Our GNN-MDAR also operates within the framework of VI using the Kullback–Leibler (KL) divergence. It is worth noting that, while VI theory is commonly applied in the distribution matching community, it is rarely explored in the context of graph-structured data with multi-source domain adaptation for regression tasks.

2.2. Domain Adaptation for Regression

In recent years, there has been an increasing amount of research on applying domain adaptation techniques to regression tasks. One category of algorithms focuses on importance weighting [31,32], which estimates the weights of training samples to address the differences in probability distributions between the source and target domains. However, the computation of these weights mostly relies on heuristic measures, which place higher demands on the initial conditions of the task and may result in suboptimal accuracy in regression predictions. Another category explores boosting-based algorithms for domain adaptation in regression tasks [33,34]. Compared to heuristic-based algorithms, the latter demonstrates better generalization performance and stability. Therefore, in our optimization process for solving spatial feature coefficients of cross-domain grid nodes, we utilize a boosting-based algorithm and draw inspiration from [35] to construct a semi-supervised loss function suitable for MDA tasks, aiming to enhance the regression prediction performance.

3. Preliminaries

To structure the REM data graph, it is necessary to perform gridification on the spatial domain of the map, as shown in Figure 1, where each grid represents a spatial area of $c \times c$. The value of c is chosen according to the spatial resolution requirements of the REM. The top-left position of each grid is denoted as a coordinate point. In this paper, the distribution of electromagnetic signal intensity in the REM is represented by the average RSRP at each spatial location.

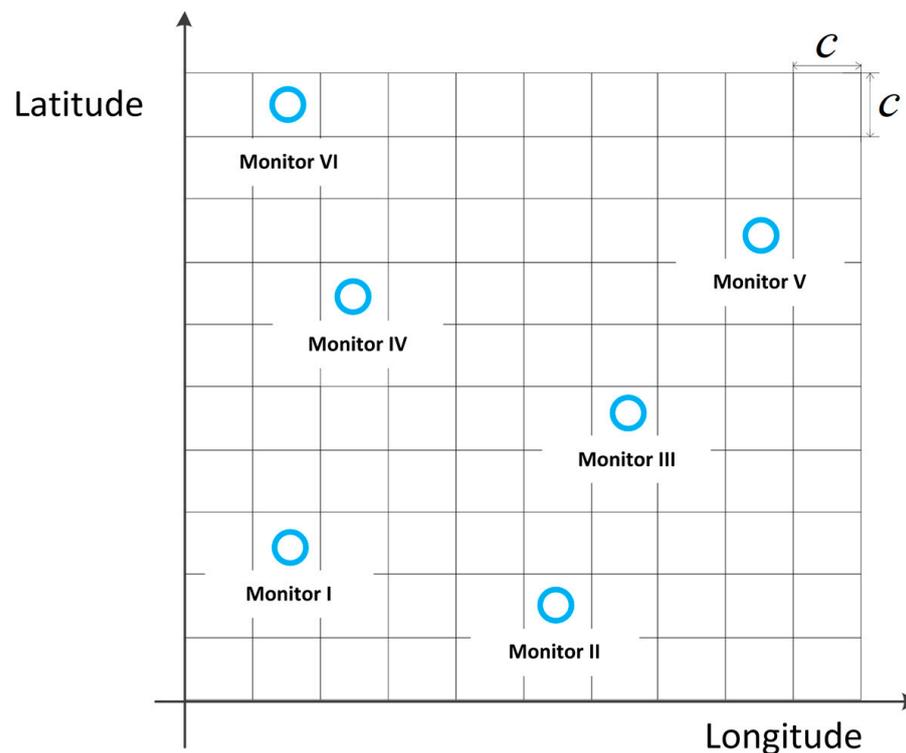


Figure 1. Grid diagram of the REM. A grid with icon O indicates the presence of monitoring stations, while an empty grid indicates the absence of monitoring stations.

3.1. Graph Structure Representation for REM

To predict the RSRP value at the grid without monitoring stations, the spatial grid is considered as a node to form a node-set $V = \{v_i\}_{i=1}^{|V|}$, and $|V|$ represents the number of grids. The original features (such as longitude and latitude coordinates, building height, altitude, and land type) of each grid are treated as node feature matrix $X \in \mathbb{R}^{|V| \times d}$, and d is the number of original features. $Y \in \mathbb{R}^{|V|}$ represents the grid RSRP vector. All grid nodes

establish edge relationships according to certain rules to generate the initial adjacency matrix $A \in \mathbb{R}^{|V| \times |V|}$. Ultimately, the graph structure G of REM can be constructed as $G = (V, A, X, Y)$.

3.2. Graph Neural Networks

Furthermore, after determining the initial graph data $G = (V, A, X, Y)$, the normalized Laplacian matrix can be calculated by $\tilde{L}_{sym} = \tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2}$, where $\tilde{A} = A + I$ represents the adjacency matrix containing self-loop and I denotes the identity matrix. The diagonal elements of the degree matrix \tilde{D} can be calculated by $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$. Then, the graph convolution [36] is performed on the REM graph structure data G by:

$$H^{(l+1)} = \sigma\left(\tilde{L}_{sym} H^{(l)} W^{(l)}\right), \quad (1)$$

where $W^{(l)}$ and $H^{(l)}$ denote the parameter matrix and node features matrix of the l th layer in GNN, respectively. Specially, $H^{(0)} = X$, and $\sigma(\cdot)$ represents the activation layer function. In this paper, we focus on two-layer graph convolutions:

$$H = GNN(X) = \sigma\left(\tilde{L}_{sym} \sigma\left(\tilde{L}_{sym} X W_0\right) W_0\right). \quad (2)$$

3.3. Problem Definition

In our work, each REM is treated as a domain. The original features of the grids in K REMs are denoted as $\mathbb{X}_S = \{X_k\}_{k=1}^K$. These grids are labeled as $\mathbb{Y}_S = \{Y_k\}_{k=1}^K$. These REMs with sufficient RSRP data serve as the multi-source domains denoted as $\mathbb{G}_S = \{G_k\}_{k=1}^K$, where $G_k = (V_k, X_k, A_k, Y_k)$, $X_k = (x_1^{(k)}, x_2^{(k)}, \dots, x_{|V_k|}^{(k)})$, and $Y_k = (y_1^{(k)}, y_2^{(k)}, \dots, y_{|V_k|}^{(k)})$. Meanwhile, the target domain is provided with limited RSRP data of REM, denoted as $G_T = (V_T, X_T, A_T, Y_T)$, where $X_T = X_T^L \cup X_T^U$ and $Y_T = Y_T^L \cup Y_T^U$. Specifically, X_T^L and X_T^U are the grids' feature matrices of the target REM with RSRP and without RSRP, respectively. $Y_T^L = (y_1^{(T)}, y_2^{(T)}, \dots, y_{|V_T^L|}^{(T)}) \in \mathbb{R}^{|V_T^L|}$ represents a vector composed of $|V_T^L|$ RSRP values from labeled grids, corresponding to $Y_T^U \in \mathbb{R}^{|V_T^U|}$ representing the vector of unlabeled grids to be predicted in the target REM. A_k and A_T are the initialized adjacency matrices of the grids in the source domain and the target domain, respectively.

Therefore, the goal of this study is to learn a mapping function $\mathcal{F}_\theta(\cdot)$, where θ represents the parameters of the learnable model. This function can reconstruct the target REM with only a small number of RSRP values by using multiple REMs with complete RSRP values as source domain data. This task can be expressed mathematically as:

$$Y_T^U = \mathcal{F}_\theta(X_T^U), \forall (Y_S, Y_T^L) \doteq \mathcal{F}_\theta(G_S, G_T). \quad (3)$$

4. Methods

This section provides a detailed introduction to our proposed GNN-MDAR model. Figure 2 shows the framework of the model, which mainly includes five modules: graph structure learner, latent feature distribution learner module, spatial matching module (SDM), GS alignment module through variational inference, and semi-supervised regression prediction module.

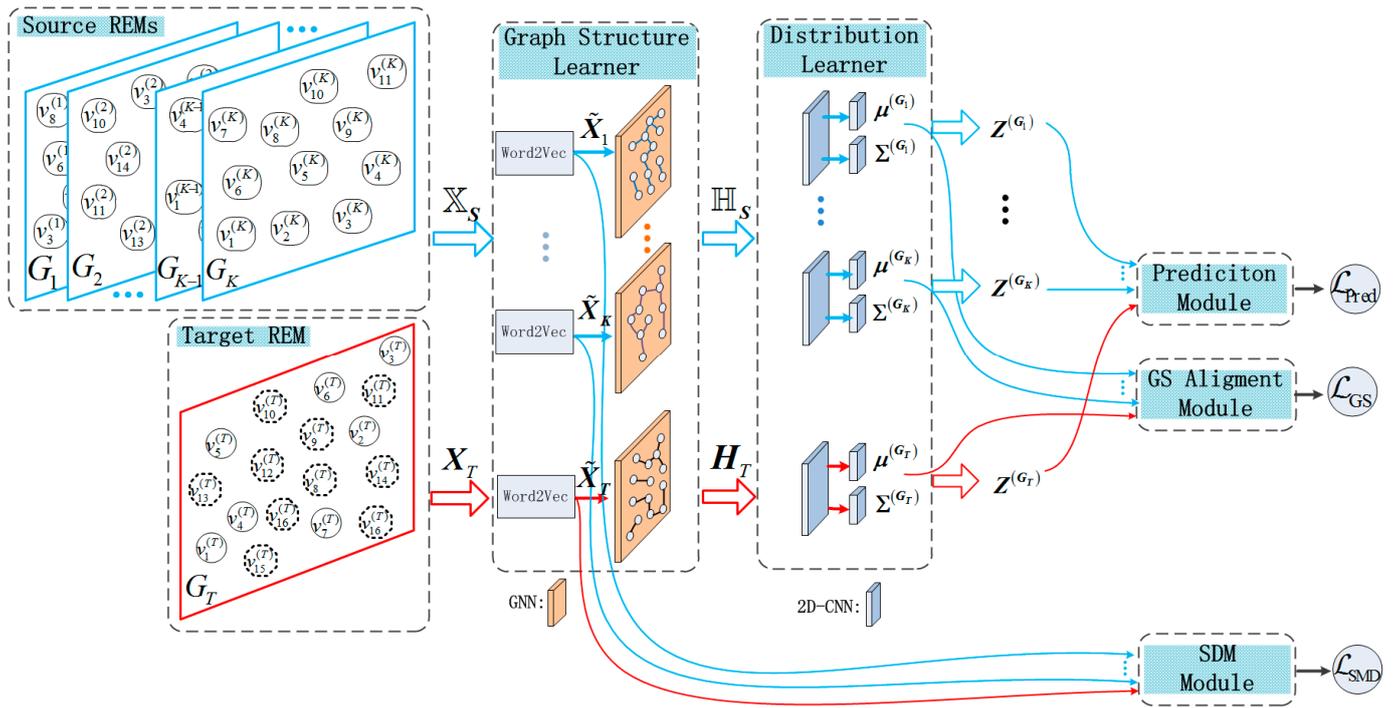


Figure 2. Framework of the proposed GNN-MDAR. In the REM on the left, solid contour nodes represent grids with RSRP label values. In the Target REM, dashed contour nodes represent grids without RSRP label values. Graph structure learner contains the Word2Vector layer and GNN layer. The distribution learner consists of Graph Convolutional Networks and 2D Convolutional Network layers, the weights of which are shared across subjects.

4.1. Graph Structure Learner

To enhance the performance of the network model, we draw inspiration from the Word2Vector concept [37] to transform the four types of original attributes of grid nodes into d_1 dimensional embedded representations $\tilde{\mathbf{X}} \in \mathbb{R}^{|\mathbf{V}| \times d_1}$:

$$\tilde{\mathbf{X}} = \text{Word2Vec}(\mathbf{X}), \quad (4)$$

where $\text{Word2Vec}(\cdot)$ is the word2Vector model, which helps to analyze and predict the propagation relationship of electromagnetic waves between grids in REM.

To initiate model learning, we use the inverse distance weights (IDW) method [38] to initialize the connection relationships between grid nodes in each REM. Based on the grid coordinate data $\{(lat_1, lon_1), \dots, (lat_{|V|}, lon_{|V|})\}$, we can establish an initialization adjacency matrix \mathbf{A} . Specially, the element $A_{i,j}$ of matrix \mathbf{A} is calculated by:

$$A_{i,j}^{(k)} = \begin{cases} 1/d_{i,j} & i \neq j \text{ and } d_{i,j} \leq th \\ 0 & \text{otherwise} \end{cases}, \quad (5)$$

where $d_{i,j} = \sqrt{((lat_i - lat_j)^2 + (lon_i - lon_j)^2)}$ denotes the spatial distance between grid v_i and grid v_j in REM. The identifier th is a threshold value for establishing initial edges between grids.

The graph structure adjacency matrices initialized by (5) are passed to the graph convolutional layer corresponding to a specific REM. GNN maps the high-dimensional embedding $\tilde{\mathbf{X}}$ to the graph structure feature space $\mathbf{H} \in \mathbb{R}^{|\mathbf{V}| \times d_2}$ through:

$$\mathbf{H} = \text{GNN}(\text{Word2Vec}(\mathbf{X}), \tilde{\mathbf{A}}), \quad (6)$$

where \tilde{A} is the optimized graph structure adjacency matrix obtained through the graph structure learner. $\tilde{A} = A$ while the model starts training. The graph structure feature of the source and target domain grids are denoted as $\mathbb{H}_S = \{\mathbf{H}_k\}_{k=1}^K$ and \mathbf{H}_T , respectively.

4.2. Distribution Learner of Latent Feature

After the GNN layer, the generated graph structure spatial features are inputted into the Distribution Learner, as shown in Figure 2. The first 2D convolutional layer of the Distribution Learner is utilized for graph structure feature fusion, and the last two 2D convolutional layers are dedicated to obtaining the approximate posterior $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ of the latent feature matrix $\mathbf{Z} \in \mathbb{R}^{|V| \times d_3}$ for REM. We denote the distributed learner constructed by a two-layer 2D CNN as $f_\phi(\cdot)$. We can calculate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ by:

$$\begin{aligned} \boldsymbol{\mu} &= f_\phi^\mu(\mathbf{H}) \\ \boldsymbol{\Sigma} &= f_\phi^\Sigma(\mathbf{H}) \end{aligned} \quad (7)$$

where $\boldsymbol{\mu} \in \mathbb{R}^{d_3}$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{d_3 \times d_3}$ are the mean vector and diagonal covariance matrix obtained from the latent features \mathbf{Z} . Therefore, we designed the graph structure learner and distribution learner for domain-adaptive variational inference of multiple REM graph structures.

4.3. Graph Structure Alignment

In the previous discussion, we mentioned that if the graph structure is highly reliable, GNN can ensure the effectiveness of downstream tasks. However, when the graph structure cannot be accurately obtained, the performance of GNN will significantly deteriorate. Therefore, it is crucial to acquire an effective graph structure. Reducing the spatial structural differences between source and target REM grids is a key issue in GNN-based MDA learning tasks [39].

Inspired by [40], we propose a multi-source adaptive graph structure alignment learning framework guided by the ‘‘Variational Information Bottleneck (VIB)’’ principle. This framework aims to learn the intrinsic relationships of electromagnetic wave propagation across different grid spaces and utilizes the shared, domain-invariant information and knowledge among multiple source REMs to guide the learning of the target domain REM prediction model. Specifically, a set of wireless REM data is divided into source and target domains. The grid feature \mathbb{X} and corresponding RSRP label values \mathbb{Y} are denoted as $\mathbb{X} = \{\mathbb{X}_S, \mathbb{X}_T\}$ and $\mathbb{Y} = \{\mathbb{Y}_S, \mathbb{Y}_T\}$, respectively. The initialization adjacency matrix for each REM is represented as $\mathbb{A} = \{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_K, \mathbf{A}_T\}$.

Assuming the optimized features set and adjacency matrices set are denoted as $\tilde{\mathbb{X}}$ and $\tilde{\mathbb{A}}$, we utilize the information bottleneck principle to compress the graph structure differences between different source and target REMs. The goal is to learn a minimal and sufficient node-level latent feature space $\mathbb{Z} = f(\tilde{\mathbb{X}}, \tilde{\mathbb{A}})$, represented as:

$$\mathbb{Z} = \underset{\mathbb{Z}}{\operatorname{argmin}} - I(\mathbf{FC}(\mathbb{Z}), \mathbb{Y}) + \beta I(\mathbb{Z}, (\mathbb{X}, \mathbb{A})), \quad (8)$$

where $I(\cdot, \cdot)$ is the mutual information used to measure the dependence between two random variables, and $\mathbf{FC}(\cdot)$ is the fully connected layer. The function $f(\cdot)$ is given by Equation (14). The first term of (8) is the prediction term, which minimizes the mutual information between the predicted values obtained through latent features and the ground truth values. In Section 4.4, we designed the training loss function for this term. The second term represents the compressed graph structure information of the cross-domain. It involves discarding the RSRP-irrelevant data in the REM grid features \mathbb{X} . β is the Lagrange multiplier that balances sufficiency and minimality. As a result, the latent features \mathbb{Z} retain

only the relevant information about the propagation of radio waves between REM grids, thereby enhancing the cross-domain invariance of the radio wave propagation.

For the original features $\mathbf{X} \in \mathbb{X}$, initialized adjacency matrix $\mathbf{A} \in \mathbb{A}$, and latent features $\mathbf{Z} \in \mathbb{Z}$ of the grid nodes, we have the following based on the definition of mutual information [41]:

$$I(\mathbb{Z}, (\mathbb{X}, \mathbb{A})) = \iint p_{\omega}(\mathbf{Z}, \mathbf{X}, \mathbf{A}) \log \frac{p_{\omega}(\mathbf{Z}|\mathbf{X}, \mathbf{A})}{p(\mathbf{Z})} d\mathbf{Z}d\mathbf{X}. \quad (9)$$

It is difficult and intractable to calculate the edge probability distribution $p(\mathbf{Z})$ practically. We use $q(\mathbf{Z}) \sim \mathcal{N}(0, \mathbf{I})$ as a variational approximation to $p(\mathbf{Z})$ and \mathbf{I} is a d_3 -dimensional identity matrix. Because of Kullback–Leiber divergence $D_{KL}(p(\mathbf{Z})||q(\mathbf{Z})) \geq 0$, we have:

$$\int p(\mathbf{Z}) \log p(\mathbf{Z}) d\mathbf{Z} \geq \int p(\mathbf{Z}) \log q(\mathbf{Z}) d\mathbf{Z}. \quad (10)$$

According to (10), the upper bound of $I(\mathbb{Z}, (\mathbb{X}, \mathbb{A}))$ can be derived as:

$$\begin{aligned} I(\mathbb{Z}, (\mathbb{X}, \mathbb{A})) &\leq \iint p_{\omega}(\mathbf{Z}, \mathbf{X}, \mathbf{A}) \log \frac{p_{\omega}(\mathbf{Z}|\mathbf{X}, \mathbf{A})}{q(\mathbf{Z})} d\mathbf{Z}d\mathbf{X} \\ &\leq \sum_{n=1}^{K+1} p_{\omega}(\mathbf{Z}_n|\mathbf{X}_n, \mathbf{A}_n) \log \frac{p_{\omega}(\mathbf{Z}_n|\mathbf{X}_n, \mathbf{A}_n)}{q(\mathbf{Z}_n)} \\ &= D_{KL}(p_{\omega}(\mathbf{Z}|\mathbf{X}, \mathbf{A})||q(\mathbf{Z})) \end{aligned} \quad (11)$$

In this way, minimizing $I(\mathbb{Z}, (\mathbb{X}, \mathbb{A}))$ can be transformed into optimizing the KL-divergence between the conditional probability distribution $p_{\omega}(\mathbf{Z}|\mathbf{X}, \mathbf{A})$ and a prior Gaussian distribution $q(\mathbf{Z})$. Therefore, the objective function of variational inference is the Evidence Lower Bound (ELBO) of Equation (11):

$$\begin{aligned} \operatorname{argmin}_{\mathbb{Z}} I(\mathbb{Z}, (\mathbb{X}, \mathbb{A})) &\doteq \operatorname{argmin}_{\omega} D_{KL}(p_{\omega}(\mathbf{Z}|\mathbf{X}, \mathbf{A})||q(\mathbf{Z})) \\ &= \operatorname{argmin}_{\omega} \mathbb{E}_{\mathbf{X} \sim p_{\omega}}[-\log q(\mathbf{Z})] - H(p_{\omega}(\mathbf{Z}|\mathbf{X}, \mathbf{A})) \end{aligned} \quad (12)$$

where $p_{\omega}(\mathbf{Z}|\mathbf{X}, \mathbf{A}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Further, ELBO in (12) can be minimized by the following equation:

$$L_{KL}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = -\frac{1}{2} \left(\log f_{\phi}^{\Sigma}(\mathbf{X}, \mathbf{A}) + \mathbf{I} - f_{\phi}^{\Sigma}(\mathbf{X}, \mathbf{A}) \left(f_{\phi}^{\Sigma}(\mathbf{X}, \mathbf{A}) \right)^{\top} - f_{\phi}^{\mu}(\mathbf{X}, \mathbf{A}) \left(f_{\phi}^{\mu}(\mathbf{X}, \mathbf{A}) \right)^{\top} \right), \quad (13)$$

where $(\cdot)^{\top}$ is the transpose operation of a matrix. The learner $f_{\phi}(\cdot)$ can be optimized to fit a function with a small divergence to prior $q(\mathbf{Z})$. With the approximate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ of the distribution learner, we use the reparameterization trick [42] for latent feature generation:

$$\mathbf{Z} = \boldsymbol{\mu} + \boldsymbol{\Sigma} \odot \boldsymbol{\varepsilon}, \quad (14)$$

where $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \mathbf{I})$ is an independent Gaussian noise and \odot denotes the Hadamard product.

After minimizing mutual information, the differences in graph structures of REM data can be reduced, while reducing the spatial structural differences between different source and target domains. $\mathbf{Z}^{(G_k)} \sim \mathcal{N}(\boldsymbol{\mu}^{(G_k)}, \boldsymbol{\Sigma}^{(G_k)})$ and $\mathbf{Z}^{(G_T)} \sim \mathcal{N}(\boldsymbol{\mu}^{(G_T)}, \boldsymbol{\Sigma}^{(G_T)})$ are the latent features for the source REM and the target subject, respectively. The loss function of the GS alignment module in Figure 2 can be expressed as:

$$\mathcal{L}_{GS} = L_{KL}(\boldsymbol{\mu}^{(G_T)}, \boldsymbol{\Sigma}^{(G_T)}) + \sum_{k=1}^K L_{KL}(\boldsymbol{\mu}^{(G_k)}, \boldsymbol{\Sigma}^{(G_k)}). \quad (15)$$

4.4. Spatial Distribution Matching

The goal of domain adaptation learning algorithms is to project features from the target domain and source domains to the same feature space. To effectively reduce distributional differences across spatially separated grids, we have adopted the idea from [43] and pro-

posed a Spatial Distribution Matching (SDM) module. SDM learns cross-domain common knowledge by matching grid features from different REMs and captures adaptively the dependencies between cross-domain grid nodes. This allows the RSRP prediction model to have good generalization capabilities on the target REM.

As shown in Figure 3, grid-wise distribution matching is weighted with a normalized $\alpha^{(k)} \in \mathbb{R}^{|V_T| \times |V_k|}$, which is named the importance matrix to learn the importance of $|V_k|$ source domain REM grids relative to $|V_T|$ target domain REM ones. The loss of spatial distribution matching is formulated as:

$$L_{SDM}(G_T, G_k) = \sum_{i=1}^{|V_T|} \sum_{j=1}^{|V_k|} \alpha_{i,j}^{(k)} d(\tilde{x}_i, \tilde{x}_j^{(k)}), \quad (16)$$

where $\alpha_{i,j}^{(k)} \in \alpha^{(k)}$ represents the importance coefficient of the grid $v_j^{(k)}$ in G_k to grid v_i in G_T . According to (4), $\tilde{x}_i \in \tilde{\mathbf{X}}_T$ and $\tilde{x}_j^{(k)} \in \tilde{\mathbf{X}}_k$ are the grid feature vectors of the grid v_i and $v_j^{(k)}$, respectively. $d(\cdot, \cdot)$ denotes distribution matching distance function such as cosine distance, MMD [44] and adversarial distance [45]. We adopt the multi-kernel MMD method. The learning scheme for the parameter matrix $\alpha^{(k)}$ is executed according to the Boosting-based Importance Evaluation Algorithm proposed in [39].

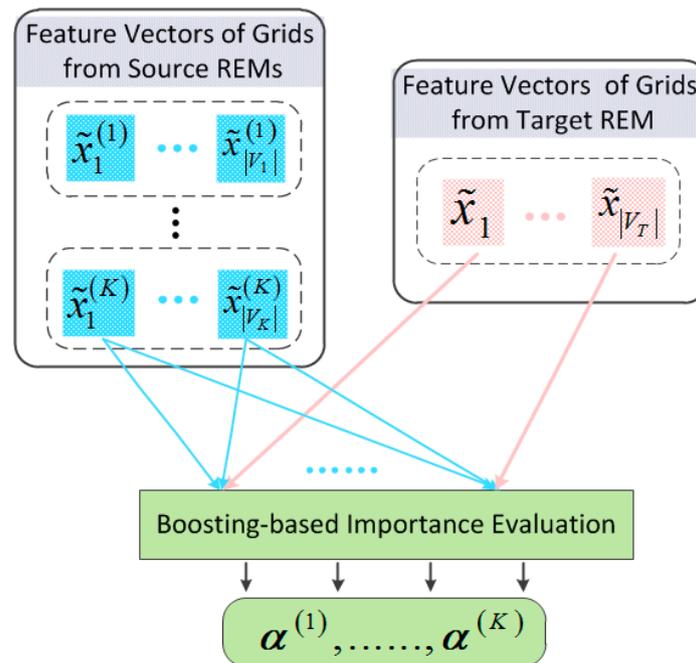


Figure 3. Spatial distribution matching module for grid nodes in cross-REMs.

4.5. Loss Function for Regression

The first term of (6) aims to train the model for accurate prediction on annotated data, which can be described by a supervised loss function:

$$L_{\text{Sup}} = \frac{1}{K} \sum_{k=1}^K \left(\frac{1}{|V_k|} \sum_{i=1}^{|V_k|} (y_i^{(k)} - \hat{y}_i^{(k)})^2 \right) + \frac{1}{|V_T^L|} \sum_{i=1}^{|V_T^L|} (y_i^{(T)} - \hat{y}_i^{(T)})^2, \quad (17)$$

where $y_i^{(\cdot)} \in \{\mathbb{Y}_S, \mathbb{Y}_T^L\}$ and $\hat{y}_i^{(\cdot)} = \mathcal{F}_\theta(x_i)$ is the predicted RSRP value of the GNN-MDAR model on the grid v_i with original features $x_i \in \{\mathbb{X}_S, \mathbb{X}_T^L\}$.

To maximize the utilization of unlabeled grid data from the target domain REM and enhance the model's robustness, we introduced a semi-supervised loss constraint term

inspired by the [31] to the supervised loss function. Specifically, we treated the grid feature data from all REMs as a whole, denoted as $\mathbb{X} = \mathbb{X}_S \cup \mathbb{X}_T$, and computed the latent feature vectors $z_i \in \mathbb{Z}$ for each grid in all REMs using Equation (12). The semi-supervised loss function can be expressed as:

$$L_{\text{SemiSup}} = \frac{1}{2} \sum_{x_i, x_j \in \mathbb{X}} w_{ij} (\mathcal{F}_\theta(x_i) - \mathcal{F}_\theta(x_j))^2, \quad (18)$$

where $w_{ij} = \exp\left(\frac{-\|z_i - z_j\|^2}{2\sigma^2}\right)$. The identifiers x_i and x_j denote the original feature vector of the grid v_i and v_j , respectively. Therefore, the first term of (6) can be described by the following loss function, which we define as the regression prediction loss function L_{pred} :

$$L_{\text{Pred}} = L_{\text{Sup}} + \lambda_1 L_{\text{SemiSup}}, \quad (19)$$

where λ_1 is the trade-off parameter of the semi-supervised loss part.

4.6. Overall Loss Function

According to (15)–(19), the final objective loss function of the GNN-MDAR model for RSRP prediction can be formulated as:

$$\mathcal{L} = L_{\text{Pred}} + \lambda_2 \frac{2}{K(K-1)} \sum_{k=1}^K L_{\text{SDM}}(G_T, G_k; \alpha^{(k)}) + \lambda_3 \mathcal{L}_{\text{GS}}, \quad (20)$$

where λ_2 and λ_3 are the trade-off parameters of the SDM part and graph structure alignment module. The target domain REM data G_T is input to the trained latent feature distribution learner and then passed through the trained module to generate predicted results for unlabeled grids.

5. Experiments

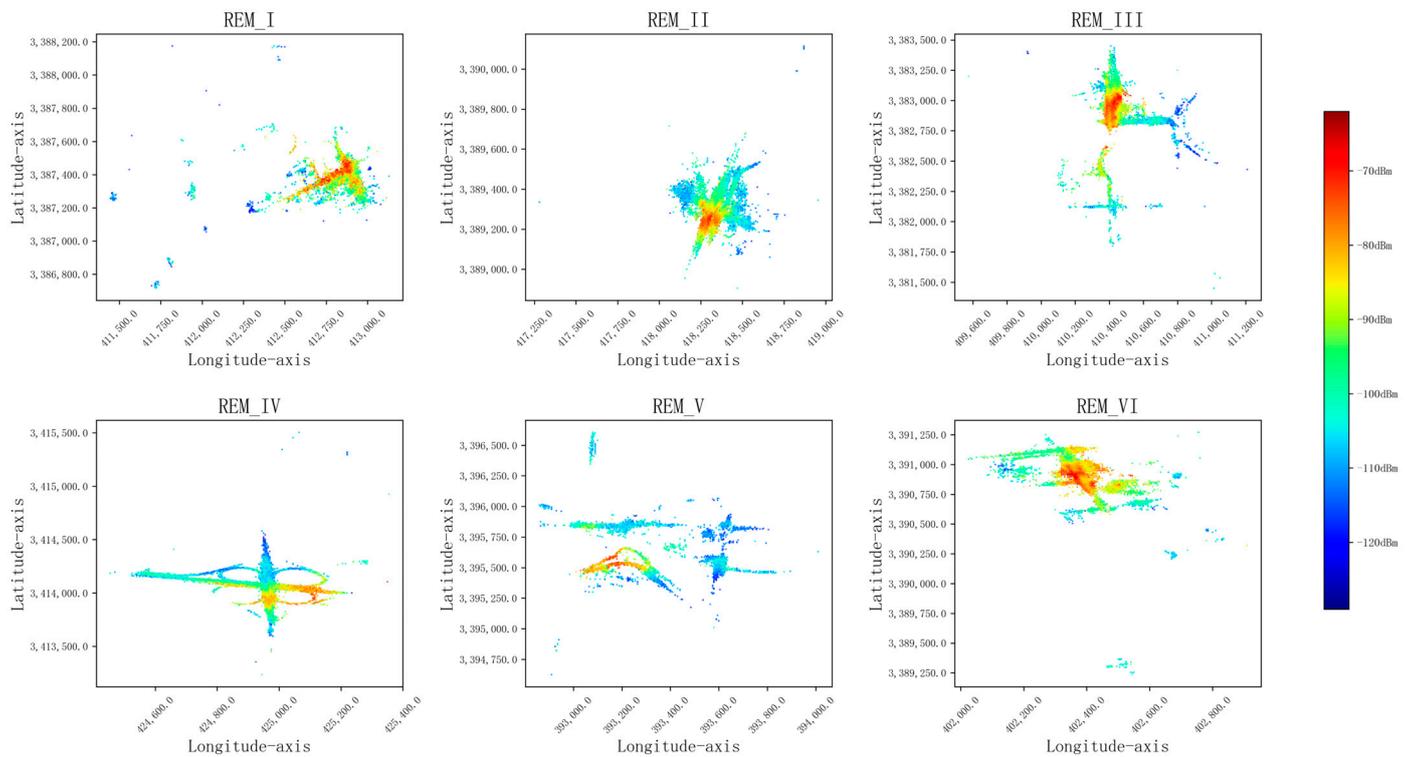
In this section, we conducted a series of numerical simulation experiments aimed at predicting grids lacking RSRP monitoring data in the target REM and evaluating the actual effectiveness of our proposed model. We first conducted an in-depth analysis of the cross-domain distribution drift characteristics of RSRP data from multiple measured areas. In addition, to ensure the transfer learning effect of shared information, we also conducted a visual analysis of the correlation between the raster feature data of the target domain and the source domain. We compared the adopted method with three other methods in terms of RSRP prediction accuracy and explored and elaborated on the experimental results using various methods.

5.1. Experiment Setup

To test the REM reconstruction performance of the proposed model and the other four models, we selected the measured RSRP data from six airspaces (numbered REM-I, REM-II, REM-III, REM-IV, REM_V, REM_VI) as the experimental dataset, with a transmitter power of 12 dBm. To construct graph-structured data, each spatial domain is rasterized according to the method in Section 3, and each grid is divided into subdomains of $5 \text{ m} \times 5 \text{ m}$ and considered as a node. The data features of each node sample are composed of longitude and latitude coordinates at the grid, building height, altitude, and clutter type information. There are 20 types of clutter (numbered 1–20), as shown in Table 1; the label of each node is the measured RSRP value at that grid. The six radio environment maps are shown in Figure 4.

Table 1. Type and index of clutter.

Index	Type of Clutter	Index	Type of Clutter
1	Oceans and Coastlines	11	High-rise Urban Buildings (40 m–60 m)
2	Lakes and Rivers	12	Middle and High-rise Buildings in Urban Areas (20 m–40 m)
3	Wetlands and Marshes	13	High-density Building Complex (<20 m) in Urban Areas
4	Suburban Open Areas	14	Multi-story Buildings (<20 m) in Urban Areas
5	Urban Open Areas	15	Low-density Industrial Building Areas
6	Roadside Open Areas	16	High-density Industrial Building Areas
7	Grasslands or Pastures	17	Suburbs
8	Shrub Vegetation	18	Developed Suburban Areas
9	Forest Vegetation	19	Rural Areas
10	Supertall Urban Buildings (>60 m)	20	CBD Commercial Zone

**Figure 4.** Radio environmental maps.

To further explore the effectiveness of the model in cross-domain data migration, especially when there are significant differences in the distribution of source domain data, we have designated the sixth REM as the target object for predicting RSRP. Based on this setting, we systematically constructed two MDA tasks. Among them, Task I focuses on knowledge transfer from the source domains I and V to the target domain VI; Task II covers the migration process from the source domains II, III, and IV to the target domain VI. Through such an experimental design, we hope to comprehensively analyze the impact of source domain data diversity on the effectiveness of cross-domain data transfer in the model. The proportion of various clutter types in six REMs is shown in Figure 5.

The key statistical distribution parameters of these six REM RSRP data are listed in Table 2, providing us with a comprehensive overview of the data characteristics. To further visually display the differences in the distribution of task data between the two domains, we have specially drawn Figure 6, which, through an intuitive graphical approach, deeply demonstrates the uniqueness of the distribution of each REM data and the comparative relationship between them.

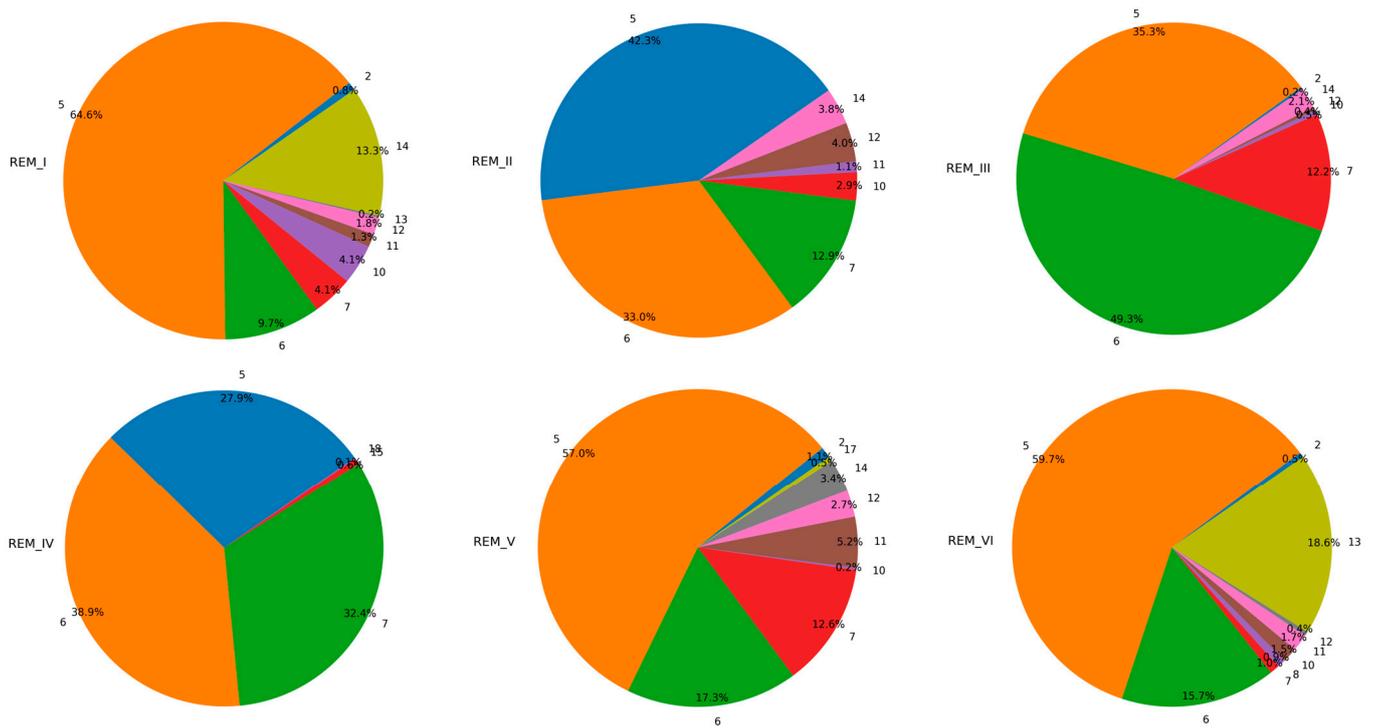


Figure 5. Clutter type distribution charts.

Table 2. Summary of the 6 REMs' data.

REM ID	Number of Grids	RSRP Statistical Distribution	
		Mean Value	Variance
I	2483	−91.21 dBm	93.74 dBm
II	3327	−90.75 dBm	90.74 dBm
III	3837	−90.97 dBm	88.89 dBm
IV	3612	−88.88 dBm	86.08 dBm
V	2312	−94.72 dBm	143.45 dBm
VI	2053	−89.53 dBm	109.78 dBm

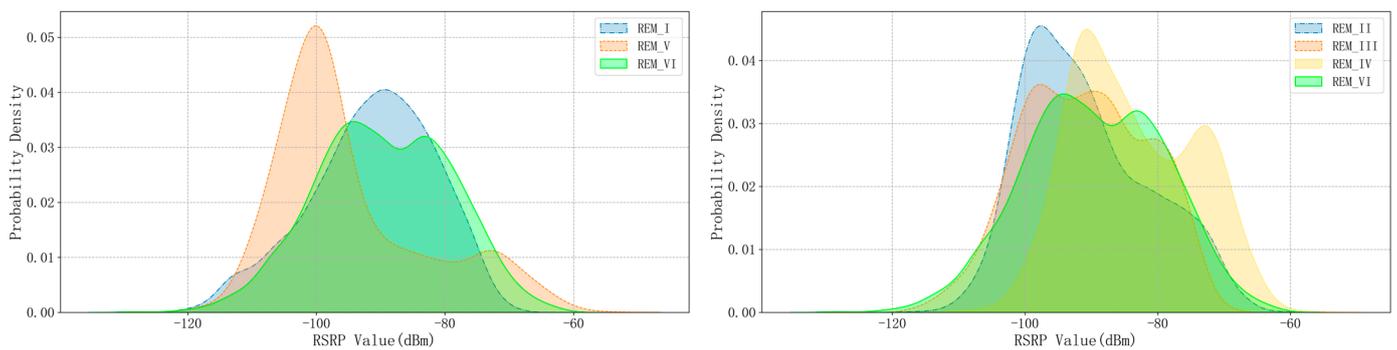


Figure 6. Probability distribution visualization of RSRP values in Task I and Task II.

To visually display and analyze the information sharing between each source domain REM grid node and the target domain REM, we selected some grid nodes on each REM (as shown in Figure 7). Figure 8 shows the correlation between these grid nodes of the five source domain REMs and the grid nodes of the target domain REM.

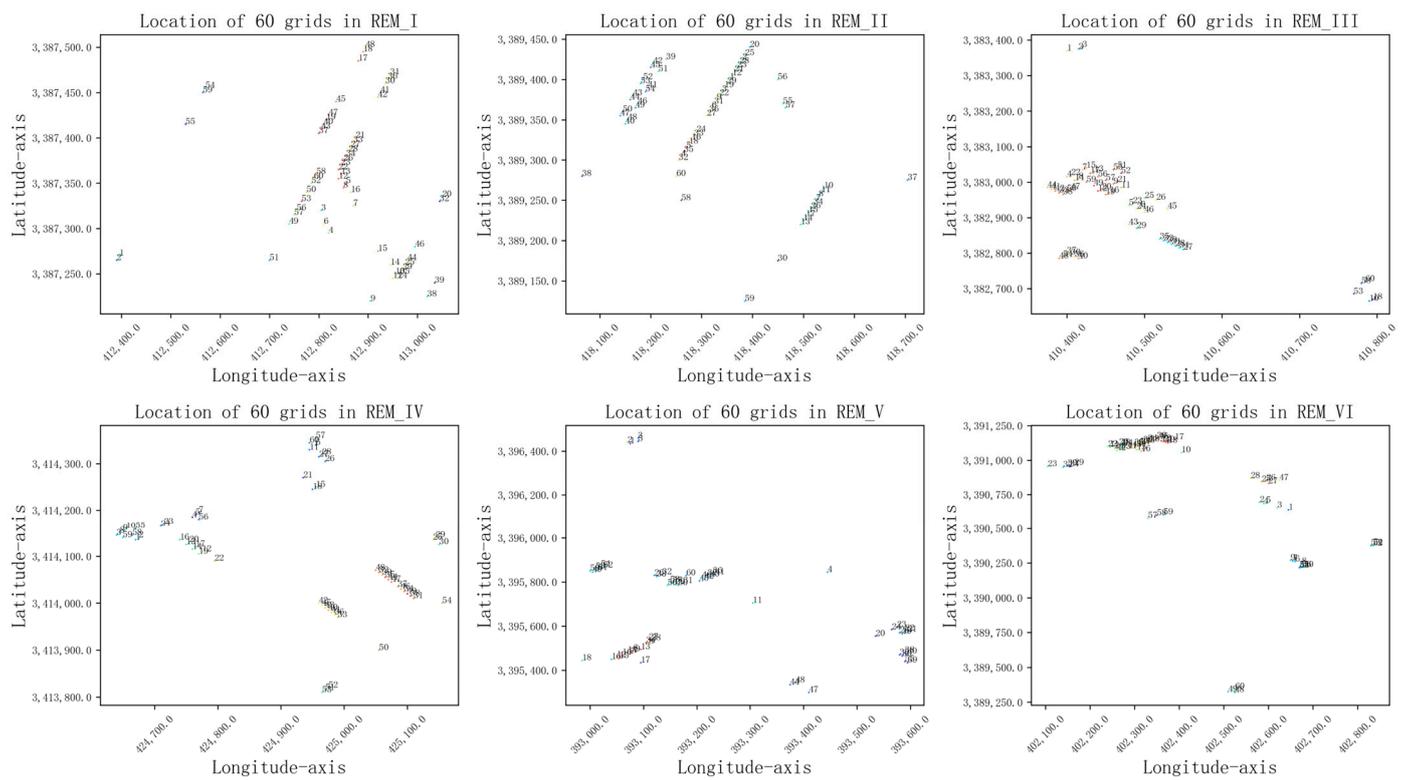


Figure 7. Visualization of 60 grids' locations selected randomly from each of the 6 REMs.

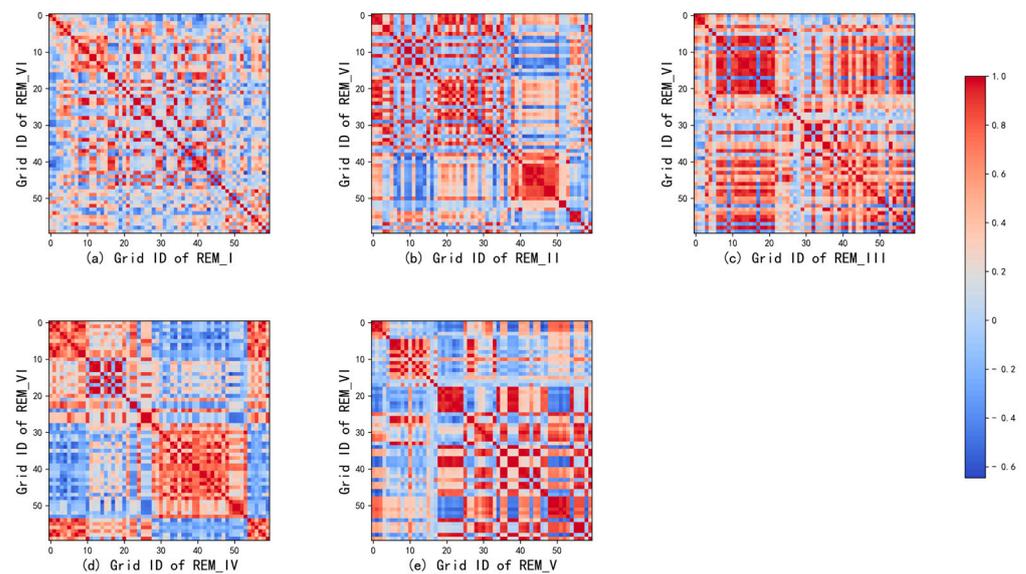


Figure 8. Visualization of feature correlation for 60 random grids between the source and target REM.

From Figure 7, we can observe the location of 60 grid nodes in REM_I, which are located in the same street as the radiation source and closely clustered near the radiation source, while the grid nodes collected in REM_VI are not located in the same street as the radiation source. The visualization results in Figure 8a also reflect the low correlation between this group of grid nodes, indicating that the grid nodes in REM_I have a weak influence on the grid nodes in the target domain REM_VI. Looking at REM_V in Figure 7, the grid nodes of No. 20–60 are located in another street far away from the radiation source and the upper left direction of the radiation source. This is similar to the spatial

characteristics of some grid nodes of No. 15–60 in the target domain REM VI. Figure 8e also shows that these columns have more high-value points than other columns.

The visualization results in Figures 7 and 8 provide us with strong evidence that there is more shared information between grids with similar node characteristics (spatial location, altitude, building height, ground object type, etc.) in the source domain and the target domain, and this shared information is crucial to the influence of radio wave propagation.

5.2. Experimental Results and Analysis

In the experiment, to generate sufficient training datasets and reduce the amount of graph structure data in each batch, we adopted a block-based training scheme to divide the REM region into multiple subgraphs by designing a lightweight GNN to achieve target domain wireless environment map prediction.

Three performance evaluation metrics commonly are used for regression prediction tasks in this paper: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE), which are defined in Equation (21):

$$\begin{aligned} MAE &= \frac{\sum_{n=1}^{|V_T^U|} |y_n^{(T)} - \hat{y}_n^{(T)}|}{|V_T^U|} \\ RMSE &= \sqrt{\frac{\sum_{n=1}^{|V_T^U|} (y_n^{(T)} - \hat{y}_n^{(T)})^2}{|V_T^U|}} \\ MAPE &= \frac{100\%}{|V_T^U|} \sum_{n=1}^{|V_T^U|} \left| \frac{y_n^{(T)} - \hat{y}_n^{(T)}}{y_n^{(T)}} \right| \end{aligned} \quad (21)$$

where $y_n^{(T)}$ and $\hat{y}_n^{(T)}$ denote prediction and truth of RSRP, respectively. By calculating the prediction results of the $|V_T^U|$ unlabeled grids according to (21), smaller values indicate higher accuracy.

We define the ratio of the number of unlabeled grids in the target REM to the total number of grids as the sampling rate $r = \frac{|V_T^U|}{|V_T|}$. In the experiment, we compared the algorithm model proposed in this paper with three benchmark models, including the GNN method that directly uses manually defined graph structure adjacency matrices [17], the CNN architecture based on block training [46], the Kriging algorithm that uses exponential semi variogram function [5], and the GAN-CRME module [13].

Next, we will discuss the model performance under two source domain combinations when $r = 20\%$.

5.2.1. Discussion on the Output Dimensions of Each Network Layer

In deep neural networks, the output dimensions of each convolutional layer profoundly affect the feature extraction ability, model capacity, and expression ability of the network model. For this purpose, we discussed the impact of the output dimensions d_1 and d_2 of the Word2Vec layer and GNN layer, as well as the dimension d_3 of the hidden feature vectors, on the model prediction error. In two domain adaptation tasks, we use the grid search method to find the optimal dimension value in {8, 16, 32, 64, 128, 256}. As shown in Figure 9, the histogram and line plots represent the MAE and RMSE values of the GNN-MDAR model for different values of d_1 , d_2 and d_3 , respectively. As the dimension values vary, the prediction error of the model will increase or decrease. The experimental results show that $d_1 = 128$, $d_2 = 64$ and $d_3 = 16$ is reasonable. As shown in Figure 9, we can see that the prediction results decrease or increase significantly in different dimensions of the node embeddings. These phenomena show that 128, 64, and 16 for the dimensions of \tilde{X} , H , Z are reasonable, respectively.

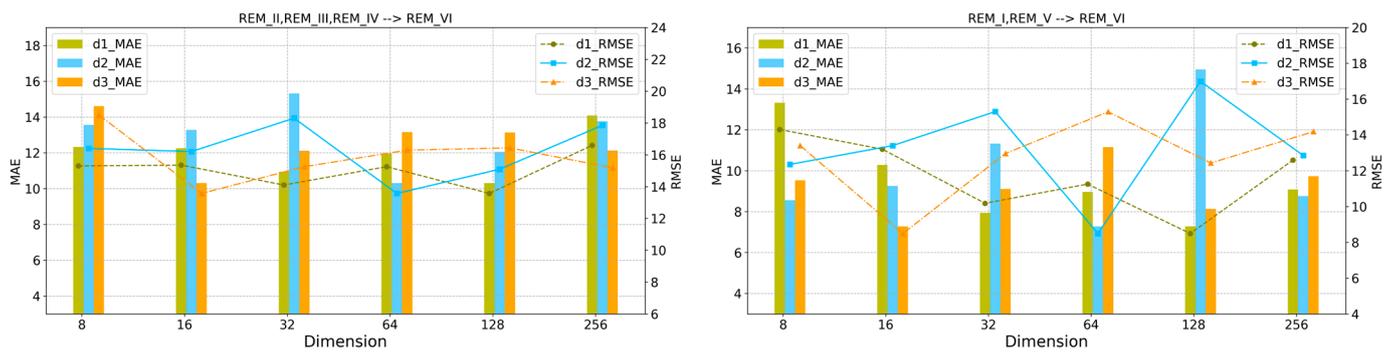


Figure 9. Effects of different output dimensions in the network layers.

5.2.2. Discussion on the Effect of Trade-Off Parameters

To further investigate the three equilibrium parameters λ_1 , λ_2 and λ_3 in (20), we also conducted comparative experiments on the impact of three factors on model performance in two domain adaptation tasks. We manually tune the balance coefficient λ_1 for semi-supervised loss term and the coefficient λ_2 of the regularization term for the difference in grid features both at $\{0.3, 0.6, 0.8, 1.2, \text{ and } 1.5\}$ and $\{1, 2, 3, 4, 5\}$, respectively. As shown in Figure 10, the reasonable setting in domain adaptation Task I is $\lambda_1 = 1.2$, $\lambda_2 = 3$, while in domain adaptation Task II $\lambda_1 = 0.6$, $\lambda_2 = 5$. The different values of λ_2 represent the differences in feature distribution between the grids of the source and target domains, meaning that larger λ_2 values will transmit more shared information for the prediction task in the target domain.

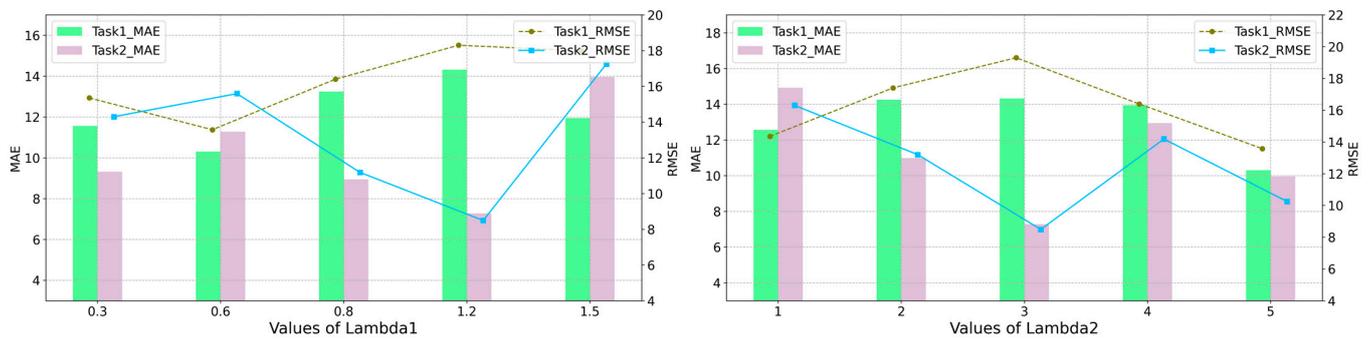


Figure 10. Effects of different trade-off parameters λ_1 and λ_2 .

The identifier λ_3 is the balance parameter that controls the alignment of cross-domain graph structures. Due to the presence of noisy data in the graph structure constructed by grids in REM during the early stages of model training, there is uncertainty in the edge relationships between grid nodes. Therefore, adopting a dynamic adjustment strategy $\lambda_3 = \frac{2}{1+e^{-10p}} - 1$ instead of fixed adaptation factors is more reasonable, where p is changed linearly from 0 to 1.

5.2.3. Discussion on the Performance of Four Prediction Models

The comparison results of the completion errors of five REM completion algorithms for the target REM_VI at a grid sampling rate of 20% are shown in Table 3. The “Datasets” column in the table represents the raster data used for training the corresponding model.

Table 3. Performance comparison of different methods for RSRP prediction of target REM_VI.

Methods	Datasets	MAE	RMSE	MAPE(%)
Kriging	REM_VI	24.45	30.11	21.15
CNN	REM_V, REM_VI	16.63	19.02	14.12
GNN	REM_V, REM_VI	15.26	18.71	13.83
	REM_I, REM_V, REM_VI	15.05	17.93	13.08
GAN-CRME	REM_I, REM_V, REM_VI	11.45	16.03	14.16
	REM_II, REM_III, REM_IV, REM_VI	8.71	12.53	10.24
GNN-MDAR	REM_I, REM_V→REM_VI	10.31	13.57	8.42
	REM_II, REM_III, REM_IV→REM_VI	7.28	8.49	6.07

In the case of limited spatial grid data and irregular distribution, the Kriging interpolation algorithm based on spatial sparsity is easy to implement, but the estimation accuracy is also the worst. Under the premise of increasing a certain amount of data, CNNs are significantly better than traditional methods. For the GNN model, we use the initialization adjacency matrix to describe the graph structure information. Under conditions of limited data volume, the predictive performance of GNN models is significantly poor. In the case of increasing training data, the performance of the GNN model did not show a particularly significant improvement due to the problem of noise in the artificially constructed graph structure. Due to the more diverse dataset distribution in Task II, the GAN-CRME exhibited a higher RMSE performance of 1.07 dB on the dataset of Task I. Additionally, as the GAN-CRME is incapable of addressing multipath fading and long-distance shadowing, its overall performance is somewhat inferior to that of GNN-MDAR.

In the GNN-MDAR framework, the spatial distribution alignment module addresses the issue of distribution skew in the dataset. Concurrently, the graph structure alignment module also plays a role in excavating more precise shared spatial graph structure features. The experimental outcomes from two domain adaptation tasks in Table 3 indicate that GNN-MDAR is capable of achieving superior REM reconstruction performance under the limited target domain data.

In the experiment, we used sampling rates of (5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 60%, 70%) for experimental analysis. As shown in Figure 11, with the increase in grid sampling rate, the number of stations containing RSRP data also increases, positively influencing the accuracy of the four methods. However, compared to deep GNN, the improvement effect of graph convolutional networks is not significant, mainly due to the noise issues present in the initialized graph structure data. As the sampling grid in the target domain increases, the data generator of the GAN-CRME can learn a more accurate representation of the target domain data distribution. Consequently, the overall predictive performance of the model is enhanced. In contrast, the performance of GNN-MDAR remains relatively stable with an increase in sampling rates. This stability is due to the GNN-MDAR model's primary reliance on the shared graph structural information from multiple source domains and the target domain. Therefore, compared to other models, the GNN-MDAR model exhibits superior overall generalization capabilities. A comparative analysis of the map reconstruction effects among five algorithms ($r = 20\%$) is presented in Figure 12. Visual inspection reveals that our proposed model exhibits superior performance compared to the other four models.

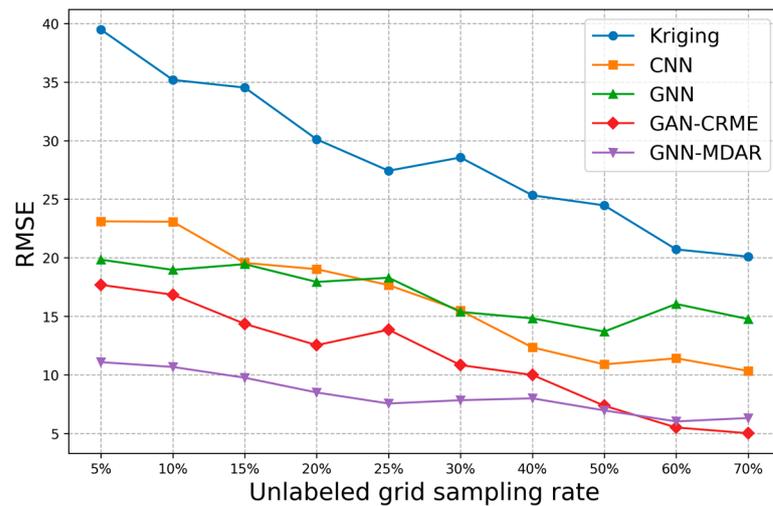


Figure 11. Performance under different sampling rates.

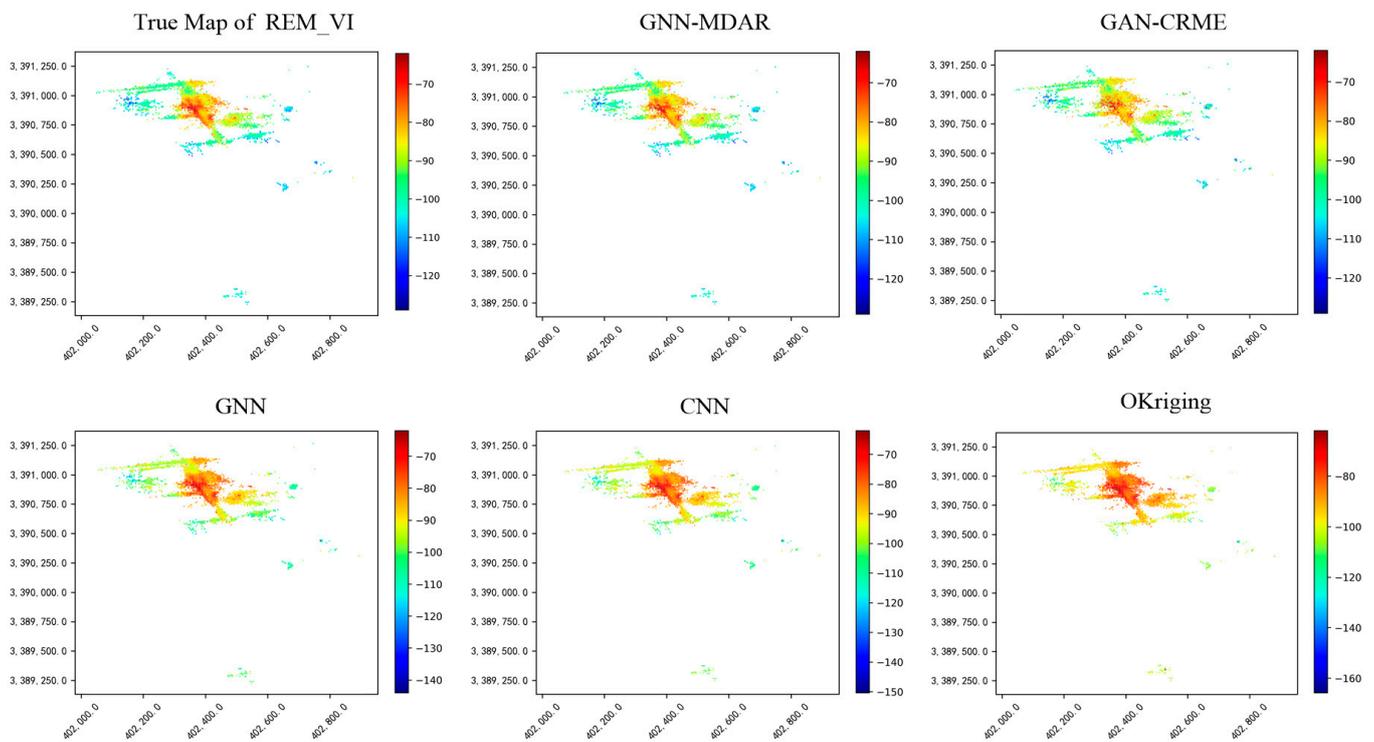


Figure 12. The true map and the maps are complemented by predicting RSRP values through four methods.

6. Conclusions and Future Works

In this paper, we propose a multi-source domain adaptive radio map construction scheme GNN-MDAR based on GNN, which aims to solve the challenge of the scarcity of spectrum data in the target REM region RSRP prediction task. This method adopts the principle of graph structure information bottleneck, realizes cross-domain graph structure alignment by variational reasoning, and ensures the compression and aggregation of data with the same structure in the latent feature space of the raster. This alignment strategy can improve the prediction performance of the multi-domain graph data distribution drift adaptive method in regression tasks. At the same time, spatial distribution matching is proposed to improve the generalization ability of the GNN-MDAR model through cross-domain grid feature distribution matching, which is helpful to transfer knowledge

from source REM with sufficient data to target REM with limited data. In addition, we develop the loss function related to the MDA REM prediction task and train the depth neural network by optimizing the loss function to solve the challenging problem of the MDA algorithm in the regression task. We use the RSRP data of six measured radio environments to compare the proposed model and the baseline model. Numerical results show that the proposed method is effective and robust to transfer learning. Specifically, through quantitative analysis, we demonstrate that our model outperforms the GAN-CRME, which achieves the best results in REM reconstruction among baseline models. When compared with GAN-based reconstruction models, our proposed model achieves an average improvement of 1.9dB in the RMSE metric under low grid sampling rates ($r = 5\text{--}25\%$) in the target domain. In future work, we are committed to the research and verification of this method in the 3D space environment and carry out further research under the conditions that the number of radio monitoring stations is more limited and the quality of spectrum data collection is worse.

Author Contributions: Conceptualization, Y.F.; methodology, X.W. and S.F.; resources, X.W. and S.F.; writing—original draft, X.W. and Y.F.; funding acquisition, S.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Basic Research Projects of the Basic Strengthening Program, grant number 2020-JCJQ-ZD-071.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author upon reasonable request.

Acknowledgments: We wish to thank Shuai Zhang for his timely assistance in analyzing a large amount of raw data samples, and for providing useful ideas for future interdisciplinary research. Zhang received his bachelor's degree from S.K.Lee Honors College, China University of Geosciences (Wuhan) in 2018 and is currently working in the East China Metallurgical Geological Prospecting Bureau of Geophysical Prospecting Team. His research interests include artificial intelligence research in spatial geographic information systems.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Miranda, R.F.; Barriquello, C.H.; Reguera, V.A.; Denardin, G.W.; Thomas, D.H.; Loose, F.; Amara, L.S. A Review of Cognitive Hybrid Radio Frequency/Visible Light Communication Systems for Wireless Sensor Networks. *Sensors* **2023**, *23*, 7815. [[CrossRef](#)]
2. Beibei, W.; Liu, K.J.R. Advances in cognitive radio networks: A survey. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 5–23. [[CrossRef](#)]
3. Üreten, S.; Yongaçoğlu, A.; Petriu, E. A comparison of interference cartography generation techniques in cognitive radio networks. In Proceedings of the 2012 IEEE International Conference on Communications (ICC), Ottawa, ON, Canada, 10–15 June 2012; pp. 1879–1883.
4. Romero, D.; Kim, S.-J. Radio Map Estimation: A data-driven approach to spectrum cartography. *IEEE Signal Process. Mag.* **2022**, *39*, 53–72. [[CrossRef](#)]
5. Boccolini, G.; Hernandez-Penalosa, G.; Beferull-Lozano, B. Wireless sensor network for Spectrum Cartography based on Kriging interpolation. In Proceedings of the 2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications—(PIMRC), Sydney, NSW, Australia, 9–12 September 2012; pp. 1565–1570.
6. Hamid, M.; Beferull-Lozano, B. Non-parametric spectrum cartography using adaptive radial basis functions. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 3599–3603.
7. Dall'Anese, E.; Bazerque, J.A.; Giannakis, G.B. Group sparse Lasso for cognitive network sensing robust to model uncertainties and outliers. *Phys. Commun.* **2012**, *5*, 161–172. [[CrossRef](#)]
8. Ding, G.; Wang, J.; Wu, Q.; Yao, Y.-D.; Song, F.; Tsiftsis, T.A. Cellular-Base-Station-Assisted Device-to-Device Communications in TV White Space. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 107–121. [[CrossRef](#)]
9. Tang, M.; Ding, G.; Wu, Q.; Xue, Z.; Tsiftsis, T.A. A Joint Tensor Completion and Prediction Scheme for Multi-Dimensional Spectrum Map Construction. *IEEE Access* **2016**, *4*, 8044–8052. [[CrossRef](#)]
10. Zhang, G.; Fu, X.; Wang, J.; Hong, M. Coupled Block-term Tensor Decomposition Based Blind Spectrum Cartography. In Proceedings of the 2019 53rd Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 3–6 November 2019.
11. Shrestha, S.; Fu, X.; Hong, M. Deep Spectrum Cartography: Completing Radio Map Tensors Using Learned Neural Models. *IEEE Trans. Signal Process.* **2022**, *70*, 1170–1184. [[CrossRef](#)]

12. Han, X.; Xue, L.; Shao, F.; Xu, Y. A Power Spectrum Maps Estimation Algorithm Based on Generative Adversarial Networks for Underlay Cognitive Radio Networks. *Sensors* **2020**, *20*, 311. [[CrossRef](#)]
13. Zhang, Z.; Zhu, G.; Chen, J.; Cui, S. Fast and Accurate Cooperative Radio Map Estimation Enabled by GAN. *arXiv* **2024**, arXiv:2402.02729.
14. Levie, R.; Yapar, Ç.; Kutyniok, G.; Caire, G. RadioUNet: Fast Radio Map Estimation with Convolutional Neural Networks. *IEEE Trans. Wirel. Commun.* **2021**, *20*, 4001–4015. [[CrossRef](#)]
15. Teganya, Y.; Romero, D. Deep Completion Autoencoders for Radio Map Estimation. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 1710–1724. [[CrossRef](#)]
16. Shen, Y.; Zhang, J.; Song, S.H.; Letaief, K.B. Graph Neural Networks for Wireless Communications: From Theory to Practice. *IEEE Trans. Wirel. Commun.* **2023**, *22*, 3554–3569. [[CrossRef](#)]
17. Lee, M.; Yu, G.; Dai, H.; Li, G.Y. Graph Neural Networks Meet Wireless Communications: Motivation, Applications, and Future Directions. *IEEE Wirel. Commun.* **2022**, *29*, 12–19. [[CrossRef](#)]
18. Eisen, M.; Ribeiro, A. Large Scale Wireless Power Allocation with Graph Neural Networks. In Proceedings of the 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Cannes, France, 2–5 July 2019; pp. 1–5.
19. Naderializadeh, N.; Eisen, M.; Ribeiro, A. Wireless power control via counterfactual optimization of graph neural networks. In Proceedings of the IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Atlanta, GA, USA, 26–29 May 2020; pp. 1–5.
20. Zhao, S.; Jiang, X.; Jacobson, G.; Jana, R. Cellular network traffic prediction incorporating handover: A graph convolutional approach. In Proceedings of the 17th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON), Como, Italy, 22–25 June 2020; pp. 1–9.
21. Chen, G.; Liu, Y.; Zhang, T.; Zhang, J.; Guo, X.; Yang, J. A Graph Neural Network Based Radio Map Construction Method for Urban Environment. *IEEE Commun. Lett.* **2023**, *27*, 1327–1331. [[CrossRef](#)]
22. Bufort, A.Y.; Leboq, L.; Cathabard, S. Data-Driven Radio Propagation Modeling using Graph Neural Networks. *TechRxiv* **2023**. [[CrossRef](#)]
23. Saenko, K.; Kulis, B.; Fritz, M.; Darrell, T. Adapting Visual Category Models to New Domains. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6314, pp. 213–226.
24. Jiang, J.; Ji, Y.; Wang, X.; Liu, Y.; Wang, J.; Long, M. Regressive Domain Adaptation for Unsupervised Keypoint Detection. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 19–25 June 2021; pp. 6776–6785.
25. Hoffman, M.; Blei, D.M.; Wang, C.; Paisley, J. Stochastic Variational Inference. *Comput. Sci.* **2012**, *14*, 1303–1347.
26. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; Volume 2, pp. 2672–2680.
27. Qu, H.; Xu, L.; Cai, Y.; Foo, L.G.; Liu, J. Heatmap Distribution Matching for Human Pose Estimation. *arXiv* **2022**, arXiv:2210.00740.
28. Zhang, C.; Cai, Y.; Lin, G.; Shen, C. DeepEMD: Few-Shot Image Classification With Differentiable Earth Mover’s Distance and Structured Classifiers. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 12200–12210.
29. Peng, H.; Sun, M.; Li, P. Optimal Transport for Long-Tailed Recognition with Learnable Cost Matrix. In Proceedings of the International Conference on Learning Representations, Virtual, 25–29 April 2022.
30. Schultze, S.; Vernaza, P.; Choi, W.; Chandraker, M. Deep Network Flow for Multi-object Tracking. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2730–2739.
31. de Mathelin, A.; Richard, G.; Deheeger, F.; Mougeot, M.; Vayatis, N. Adversarial Weighting for Domain Adaptation in Regression. In Proceedings of the 2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI), Washington, DC, USA, 1–3 November 2021; pp. 49–56.
32. Chen, X.; Wang, S.; Wang, J.; Long, M. Representation Subspace Distance for Domain Adaptation Regression. In Proceedings of the 38th International Conference on Machine Learning, PMLR, Virtual, 18–24 July 2021; Volume 139, pp. 1749–1759.
33. Kutbi, M.; Peng, K.-C.; Wu, Z. Zero-Shot Deep Domain Adaptation with Common Representation Learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 3909–3924. [[CrossRef](#)] [[PubMed](#)]
34. Wang, B.; Mendez, J.A.; Cai, M.B.; Eaton, E. Transfer learning via minimizing the performance gap between domains. In Proceedings of the 33rd International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; pp. 10645–10655.
35. Singh, A.; Chakraborty, S. Deep Domain Adaptation for Regression. In *Development and Analysis of Deep Learning Architectures*; Pedrycz, W., Chen, S.-M., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 91–115.
36. NT, H.; Maehara, T.; Murata, T. Revisiting Graph Neural Networks: Graph Filtering Perspective. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 8376–8383.
37. Zhou, Z.; Meng, L.; Tang, C.; Zhao, Y.; Chen, W. Visual Abstraction of Large Scale Geospatial Origin-Destination Movement Data. *IEEE Trans. Vis. Comput. Graph.* **2019**, *25*, 43–53. [[CrossRef](#)] [[PubMed](#)]

38. Wen, X.; Fang, S.; Xu, Z.; Liu, H. Joint Multidimensional Pattern for Spectrum Prediction Using GNN. *Sensors* **2023**, *23*, 8883. [[CrossRef](#)] [[PubMed](#)]
39. Chen, D.; Zhu, H.; Yang, S.; Dai, Y. Unsupervised multi-source domain adaptation with graph convolution network and multi-alignment in mixed latent space. *Signal Image Video Process.* **2023**, *17*, 855–863. [[CrossRef](#)]
40. Sun, Q.; Li, J.; Peng, H.; Wu, J.; Fu, X.; Ji, C.; Yu, P.S. Graph Structure Learning with Variational Information Bottleneck. *arXiv* **2021**, arXiv:2112.08903. [[CrossRef](#)]
41. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 1st ed.; Tsinghua University Press: Beijing, China, 2006.
42. Yoo, J.; Jeon, H.; Jung, J.; Kang, U. Accurate Node Feature Estimation with Structured Variational Graph Autoencoder. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 14–18 August 2022.
43. Du, Y.; Wang, J.; Feng, W.; Pan, S.; Qin, T.; Xu, R.; Wang, C. AdaRNN: Adaptive Learning and Forecasting of Time Series. In Proceedings of the 30th ACM International Conference on Information and Knowledge Management, New York, NY, USA, 1–5 November 2021; pp. 402–411.
44. Gretton, A.; Sriperumbudur, B.; Sejdinovic, D.; Strathmann, H.; Kenji, F. Optimal kernel choice for large-scale two-sample tests. In Proceedings of the Advances in Neural Information Processing Systems 25 (NIPS 2012), Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1205–1213.
45. Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; Marchand, M.; Lempitsky, V. Domain-Adversarial Training of Neural Networks. In *Domain Adaptation in Computer Vision Applications*; Csurka, G., Ed.; Springer International Publishing: Cham, Switzerland, 2017; pp. 189–209.
46. Hashimoto, R.; Suto, K. SICNN: Spatial Interpolation with Convolutional Neural Networks for Radio Environment Mapping. In Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), Fukuoka, Japan, 19–21 February 2020; pp. 167–170.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.