A Multi-Task Learning Model for Super Resolution of Wireless Channel Characteristics

Xiping Wang1, Zhao Zhang1, Danping He12, Ke Guan12, Zhangdui Zhong12, Dongliang Liu34, Shahid Mumtaz2 and Saba Al-Rubaye6

State Key Laboratory of Railway Traffic Control and Safety, Beijing Jiaotong University, Beijing, China1
Frontiers Science Center for Smart High-speed Railway System, Beijing, China2
State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen, Guangdong, China3
ZTE Corporation, Shenzhen, Guangdong, China4
Instituto de Telecomunicações, Aveiro, Portugal5
School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, UK6
{16211133,18211026,danpinghe,kguan,zhdzhong}@bjtu.edu.cn12, liu.dongliang1@zte.com.cn34

Abstract—Channel modeling has always been the core part in communication system design and development, especially in 5G and 6G era. Traditional approaches like stochastic channel modeling and ray-tracing (RT) based channel modeling depend heavily on measurement data or simulation, which are usually expensive and time consuming. In this paper, we propose a novel super resolution (SR) model for generating channel characteristics data. The model is based on multi-task learning (MTL) convolutional neural networks (CNN) with residual connection. Experiments demonstrate that the proposed SR model could achieve excellent SR performances in mean absolute error and standard deviation of absolute error. Advantages of the proposed model are demonstrated in comparisons with other SOTA deep learning models. Ablation study also proved the necessity of multi-task learning and techniques in model design. The contribution in this paper could be helpful in channel modeling, network optimization, positioning and other wireless channel characteristics related work by largely reducing workload of simulation or measurement.

Index Terms—Wireless channel modeling, ray-tracing (RT), super resolution (SR), multi-task learning (MTL), convolutional neural network (CNN),

I. INTRODUCTION

Thanks to the fast advances in the fifth generation (5G) wireless communication, our world is stepping into the era of Internet of Everything (IoE) [1]. To enable ultra-low latency and high reliability communication services in dense connected areas like urban districts, accurate wireless channel model is a necessity. Correct knowledge of the propagation channel is also critical for radio coverage estimation, network optimization and for many related applications [2]. Naturally, channel modeling is seen as the foundation for planning and optimizing communication and related systems [3].

Channel modeling is the process of characterizing the propagation principles of radio waves in realistic environments, and provides insightful theoretical guidance for the design, deployment and optimization of communication systems. Generally, stochastic channel modeling (SCM) and ray-tracing (RT) based deterministic modeling are two main modeling approaches [4]. For SCM, channel characteristics like path loss (PL), propagation condition (line of sight or non line of sight), delay spread, angular spreads and Rician K-factor are required to generate channel coefficients and thereafter to model the wireless channel. Massive channel measurements in different environments, which are usually time-consuming and expensive, must be conducted to obtain the necessary channel characteristics data [5]. On the other side, RT based modeling approach can generate accurate channel data if given precise propagation environment and configuration, but at the cost of high computational complexity and enormous calculation time [6]. Therefore, a fast and reliable channel characteristics data generation method will effectively address the limitations above.

One of the promising solutions is machine learning (ML) method. The recent surge of ML is revolutionizing almost every branch of science and technology, including wireless channel modeling. The wireless channel is a time-varying nonlinear system, which contains multi-dimensional information in time, spatial and frequency domain. Machine learning has very powerful learning and inferring capabilities. It can automatically learn from channel data so that the structural relationship between data in complex environments can be extracted to approximate nonlinear systems. Moreover, machine learning is very efficient in mining information from high dimensional data, which can significantly expedite data processing.

Current research in this area are still inadequate. Most of recent works focused on predicting only one of the channel characteristics like PL (Radio Coverage Prediction) [2]. The input is mostly restricted to building information, satellite image [7], map [8] and just one of the characteristics. Few works attempt to estimate several channel characteristics at the same time. As for ML methods, traditional algorithms such as Random Forests (RF), Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) as well as Deep Learning (DL) methods such as Convolutional Neural Network (CNN) [8], Transformer [9] and Generative Adversarial Network (GAN) [10] are frequently employed. Also, the estimation targets are usually restricted to one or two characteristics. In general, few works have studied Multi-Task Learning (MTL) models for generating channel characteristics data with several characteristics as input. The connection between different characteristics awaits to be exploited.

In this paper, instead of prediction or estimation, we propose a super resolution (SR) model for channel characteristics.

The support of ZTE Corporation is gratefully acknowledged.

© 2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
The model is based on MTL CNN with residual connection. Overview of our work is shown in Fig.1. Given 3D models of urban areas and the corresponding EM parameters, CloudRT platform outputs channel characteristics dataset and the dataset is used for SR model training. Data from the dataset are degraded into low resolution data as input. Original high resolution data are used as the ground truth. MTL loss is employed to better balance multi tasks. We evaluate our proposed SR model by ablation study and comparisons with other DL models. Specifically, we make the following contributions:

- In dense urban areas, RT simulation by self developed CloudRT was conducted and channel characteristics dataset are constructed based on simulation results.
- A residual network based MTL SR model is proposed. Weighted masks are added in loss function. Homoscedastic uncertainty is employed to balance the single task losses during training. Residual connection and iterative up-and-down technique are implemented in CNN blocks for better SR performances.
- Training process and results of evaluation are provided. The proposed SR approach generally performs better than other SOTA DL models. Comparing with baseline, the proposed approach could achieve very good SR results in all channel characteristics targets and deteriorate significantly less than baseline with larger scale factor. Ablation study proved that the techniques in training and model design are necessary.

II. SIMULATION AND DATASET CONSTRUCTION

This section describes the procedures of RT simulation and dataset construction. In this paper, we used self developed CloudRT [11][12] platform to generate channel characteristics data.

RT approach is widely used to generate accurate channel characteristics in a specific environment. Based on 3D electronic map and electromagnetic (EM) parameters provided by ITU-R P.1238-7, RT simulation is conducted in dense urban areas on CloudRT platform. More than 100 simulation regions are manually selected from urban areas of four major cities in China: Beijing, Shanghai, Hangzhou and Xi’an. The simulation regions are 1km × 1km squares with 3D electronic map resolution of 200 × 200. Transmitter (Tx) is located on one of the high buildings near the center of simulation region, and a large LOS area should be maintained. The receivers (Rx) are located 2 meters above ground, uniformly distributed by distance of 5 meters on the horizontal plane. Only receivers outside buildings are considered in simulation. Table I summarizes the simulation configuration and details of it can be found in [13].

$$K_p = \frac{\sum_{i=1}^{N_r} P_i - P_0}{\sum_{i=1}^{N_r} P_i}$$

(1)

The channel characteristics dataset is constructed based on RT simulation results. Definitions of path loss (PL), root mean squared (RMS) delay spread (DS), RMS azimuth ($\phi$) angular spread of arrival, RMS elevation ($\theta$) angular spread of arrival, line of sight (LOS) and non line of sight (NLOS) follows the usual. We redefined Racian-K factor as multi-path power ratio $K_p$ in (1), where $P_1$ and $P_0$ is the power of ray 1 and ray of
Super-resolution is a notion of recovering High Resolution (HR) data from the Low Resolution (LR) data [14]. The LR data mostly originates from HR data with a process of degradation:

\[ I_{LR} = D(I_{HR}; \delta) \]  

where \( I_{LR}, I_{HR}, D, \delta \) denote LR data, HR data, degradation process which is usually unknown, and parameters of degradation such as scaling factor. Usually \( I_{LR}, I_{HR} \) are the input training data and the ground truth correspondingly.

SR can be described as the model of recovering \( I_{LR} \) to \( I_{HR} \). The recovery (approximation) result is denoted as \( \hat{I}_{HR} \). \( \mathcal{F} \) represents the recovery model with parameters \( \theta \). Take DL as the example, \( \mathcal{F} \) is the neural network while \( \theta \) is the weights and other related parameters in the neural network.

\[ \hat{I}_{HR} = \mathcal{F}(I_{LR}; \theta) \]  

To this end, the objective of SR is as follows:

\[ \hat{\theta} = \arg \min_{\theta} \mathcal{L}(\hat{I}_{HR}, I_{HR}) + \lambda \Phi(\theta) \]  

where \( \mathcal{L} \) is the loss function between recovery result \( \hat{I}_{HR} \) and ground truth \( I_{HR} \). \( \Phi \) is the regularization term with parameters \( \theta \) and \( \lambda \) is the weight factor. The best SR model would minimize the loss to the least with respect to \( \theta \).

### B. Residual Network based Multi-Task Learning Model

The proposed residual network based MTL model consists of two parts: backbone part and fine-tune part, as in Fig. 2. The backbone part is to extract high-dimensional features from input data while the fine-tune part focuses on each of the tasks to achieve the best SR performances.

The CNN block contains two convolutional layers with activation function ReLU. The size of data during convolution remains \( 200 \times 200 \). The number of CNN blocks \( N \) is set to be 3 in our work, but a larger number of blocks may also achieve excellent SR performances based on our previous experiments. To be noted, the numbers of channels in convolutional layers are set to be iterative up-and-down, as shown in the illustration of CNN block. The iterative up-and-down technique can filter irrelevant information in input data. Residual connection is also necessary for good SR performances. Several CNN blocks are concatenated with residual connection to construct the backbone part. In fine-tune part, 6 lightweight models are designed for the corresponding targets. Lightweight models are two layer CNN like in CNN block but number of channels remain fixed. We also observed a little bit better performances of SR with other models in fine-tune part. Note that for NaN/LOS/NLOS, outputs of the fine-tune nets are the probabilities of corresponding propagation conditions.

### C. Loss Functions and Evaluation Metrics

As explained in Section II, only receivers outside the building are considered in simulation. As a result, NaN values are given real number to guarantee that the input data is a regular matrix or tensor. Also, there will always be some...
ground truth values (elements) in the matrix of input data. However, both NaN values and ground truth values are trivial in SR process. To keep the loss function continuous and to help ML models concentrate on SR values of desired positions, weighted loss is introduced as the mask of NaN values $M_{na}$ and the mask of ground truth $M_{gt}$, which are illustrated in Fig. 3.

Several kinds of loss functions and evaluation metrics such as **pixel loss**, and **content loss** are widely used in image SR tasks. Differently, only **pixel loss** is accepted in channel characteristics SR task. According to previous experiments, L1 (norm) loss is employed for training as it performs better than L2 loss and peak signal noise ratio (PSNR).

During training, the SR process is essentially to classify between LOS and NLOS area. For the rest targets, the SR process is regression. Due to the difference between classification and regression, the loss functions are categorized into two scenarios where L1 norm are used for regression and cross entropy are used for classification, as in (7) and (8). The loss function $L_m$ in training can be described as follows:

$$L_m(\hat{I}, I) = \begin{cases} \text{loss}_{cr}(\hat{I}, I), & \text{if target } m \text{ is LOS/NLOS} \\ \text{loss}_1(\hat{I}, I), & \text{others} \end{cases}$$

(5)

where $\hat{I}$ and $I$ are SR recovered data and ground truth. Before the calculation of loss function, both $\hat{I}$ and $I$ should be weighted by **Hadamard Product** with $M_{gt}$ and $M_{na}$:

$$I_{weighted} = I_{original} \circ M_{na} \circ M_{gt}$$

(6)

$n$, $h$, $w$, represents the number of values to be estimated (SR) and side lengths of input data. $k$ stands for the class among LOS/NLOS/NaN in (8).

$$\text{loss}_1(\hat{I}, I) = \frac{n}{(hw)^2} \sum_{i,j} |\hat{I}_{i,j} - I_{i,j}|$$

(7)

$$\text{loss}_{cr}(\hat{I}, I) = \frac{n}{(hw)^2} \sum_{i,j,k} I_{i,j,k} \log \hat{I}_{i,j,k}$$

(8)

During the pre-train stage, homoscedastic uncertainty [15] is employed to balance the single-task losses as in (9). Not only the weights in neural networks $W$ but also noise parameters $\sigma$ are trainable and updated through standard back propagation during training.

$$L_{MTL}(W, \sigma) = \sum_m L_m \frac{1}{2\sigma_m^2} + \sum_m \log(\sigma_m)$$

(9)

The evaluation metric is much simpler than training loss. Basically, values of receiver located in building areas (NaN values) and ground truth values are not included in calculation. The metric for evaluating SR performance of LOS/NLOS is **classification accuracy**. For the rest targets, **mean absolute error (MAE)** and standard deviation of absolute error (SDE) are regarded as the metrics.

### IV. Experiments

#### A. Training Configuration and Implementation

In this work, DL training are performed by PyTorch 1.10.2 on a work station with 1 NVIDIA GeForce RTX 3090 GPU, Intel Core i9-9900K CPU and 32 GB DDR4 RAM.

The training process are divided into two stage: **pre-train stage** and **fine-tune stage**. In **pre-train stage**, (9) is used as the loss function for back propagation and weights of the entire SR model are updated. Next, in **fine-tune stage**, only weights of **fine-tune part** in the SR model are updated while **backbone part** remains unchanged. The purpose of this design is to obtain a general feature extractor (**backbone part**) for channel characteristics SR. By doing so, it’s very fast and convenient to concatenate lightweight models for specific tasks when comparing with DL models aiming at only one channel characteristic.

Data augmentation is employed to the training set for enhancing diversity of the inputs. The training set is transformed by rotation of 90, 180 and 270 degree as well as horizontally and vertically flipping. As a result, the training dataset is incremented by 5 times.

The proposed SR model is trained for 100 epochs in both of two stages. The learning rate is set as 0.00001. Adam optimizer is used for gradient descent. Batchsize is set as 1. The channel characteristics dataset is randomly split into training set and test set by ratio of 7:3. Experiments with scale factor 2, 4 and 8 were conducted.

#### B. Backbone Part

As mentioned above, **backbone part** of the proposed SR model is regarded as a general feature extractor. The comparison of DL models is summarized in Table III. Our proposed SR model could achieve 3.26 dB for MAE of PL with scale factor 2. Under the same configuration, several kinds of DL models including ResNet50, vision transformer (ViT) [16] and GAN-SR [17] were tested. Performances of these SOTA DL models are not satisfying. For **backbone part**, ResNet50 is the best in tested DL models but the MAE is around 7-8 dB. After several experiments of ViT by changing the number of transformer encoder and the number of heads in attention layer, the best result of MAE is around 8 dB. Performances of GAN are much worse than CNN and ViT. The best results of GAN is higher than 12 dB for MAE. Moreover, few experiments show that the loss of generator and the output of discriminator (probability that $\hat{I}_{HR}$ is real) converge at the end. Nash equilibrium between generator and discriminator could hardly be reached. For other targets, SR results of these DL models are also much worse than those of the proposed model.

The reasons of why SOTA DL models doesn’t work well are analyzed. First, the data size of channel characteristics dataset is much smaller than popular computer vision datasets like CIFAR-10, ImageNet and MINST. Second, texture, style and smoothness of images should be considered in image SR but not regarded as objectives in characteristics SR task. As a result, deeper models with larger number of parameters are more...
liable to over-fitting. The comparison of SR performances and model complexity are presented in Table III. For the proposed model, both the number of parameters (Params) and floating point of operations (FLOPs) are much smaller than other DL models.

|                  | Proposed Model | ResNet50 | ViT | GANSR |
|------------------|----------------|----------|-----|-------|
| **Params / k**   | 4.25           | 8610     | 13040 | 28670 |
| **FLOPs / GMac** | 0.34           | 2.58     | 10.33 | 12.42 |
| **MAE of PL / dB** | 3.26 (best)   | 7-8      | 7-8  | >12   |

**TABLE III: Comparison of Deep Learning Models**

C. Performance of Proposed Model

MAE and classification accuracy of 6 SR targets during training process are shown in Fig. 4. Clearly for six targets, the MAE and classification accuracy of SR are decreasing during training and converge at the end. In pre-train stage (1-100 epochs), MAE and classification accuracy decrease very quickly with some minor jitters due to MTL loss. This indicates that the proposed model and MTL loss is effective as a feature extractor to achieve fairly good SR performances with multi inputs and tasks. In fine-tune stage (101-200 epochs), few jitters are observed. MAE and classification accuracy decrease slowly to converge at the end so that better SR performances could be achieved in fine-tune stage. We also observed that given scale factor as 8, the classification accuracy of LOS/NLOS jitters greatly. We will look into this part in future research.

In general, the best SR results achieved by the proposed model is demonstrated in Table IV and Table V. Compared with bilinear interpolation, both MAE and SDE of the proposed model is far smaller. Moreover, SR performances of the proposed model deteriorate significantly less than bilinear interpolation when scale factor is relatively larger (4 or 8).

D. Ablation Study

Ablation study was conducted to investigate the effectiveness of MTL and techniques in the proposed model. MTL represents the proposed model without residual connection and iterative up-and-down technique. It is set as the baseline model for comparison. STL means the baseline model with only single task (PL) during training.

Table VI demonstrates the cumulative performance gain of MAE and SDE in ablation study. The evaluation metric is PL. Apparently, SR of single task learning is incomparable to MTL. RES Residual connection and iterative up-and-down technique could reduce MAE and SDE for 20% on average. DA could also enhance SR performance in some extent. It’s noticed that the enhancement of DA and RES decreases much when scale factor are large (4 and 8).

V. CONCLUSION

In this paper, a novel residual network based MTL model is proposed for SR of wireless channel characteristics. RT simulation was conducted and channel characteristics dataset were constructed based on simulation results. Weighted masks are introduced in loss function which can help better fit randomly
distributed building regions. A general MTL model with two stage training methods are proposed. The proposed model could achieve SR results of PL with MAE of 3.26 dB and 98% classification accuracy of LOS/NLOS areas given scale factor as 2. It also outperforms other SOTA DL models and the reasons are discussed. The proposed model demonstrates huge advantages in channel characteristics SR tasks especially when scale factor is relatively large. Ablation study also proved the necessity of residual connection, iterative up-and-down technique and multi-task learning. In future, we will continue the study of channel characteristics SR problem by refining proposed MTL model on network structure and MTL loss function. Relations between upper bounds of SR performances and scale factor will also be explored.

ACKNOWLEDGMENT

This work is supported by National Key RD Program of China under Grant 2020YFB1806604, NSFC under Grant 61901029 and Grant U1834210, Beijing Natural Science Foundation L212029, ZTE Corporation and State Key Laboratory of Mobile Network and Mobile Multimedia Technology. We appreciate Haoyang Zhang and Wenbin Wang for their help in RT simulation and dataset generation.

REFERENCES

[1] J. CAO, L. XU, R. Abdallah, and W. SHI, “An os for internet of everything: Early experience from a smart home prototype,” ZTE Communications, vol. 15, no. 4, pp. 12–22, 2019.
[2] S. Mohammadjafari, S. Roginsky, E. Kavurmacioglu, M. Cevik, J. Ethier, and A. B. Bener, “Machine learning-based radio coverage prediction in urban environments,” IEEE Transactions on Network and Service Management, vol. 17, no. 4, pp. 2117–2130, 2020.
[3] F. Xiao, X. Xie, Z. Li, Q. Deng, A. Liu, and L. Sun, “Wireless network optimization via physical layer information for smart cities,” IEEE Network, vol. 32, no. 4, pp. 88–93, 2018.

TABLE IV: Comparison of MAE between Proposed model and Bilinear Interpolation

| Scale | Targets | PL | $K_p$ | DS | $\phi$ | $\theta$ | LOS/NLOS |
|-------|---------|----|-------|----|-------|---------|-----------|
| 2     | Proposed| 3.26| 0.47  | 6.84| 1.6   | 5.34    | 98%       |
|       | Bilinear| 16.58| 6.73  | 13.46| 12.97 | 16.6    | 85%       |
| 4     | Proposed| 4.29| 0.69  | 9.99| 2.31  | 8.55    | 95%       |
|       | Bilinear| 26.76| 11.06 | 20.41| 21.13 | 26.53   | 72%       |
| 8     | Proposed| 5.80| 0.99  | 14.09| 3.57  | 12.53   | 87%       |
|       | Bilinear| 36.71| 15.34 | 27.9 | 29.71 | 36.92   | 65%       |

TABLE V: Comparison of SDE between Proposed model and Bilinear Interpolation

| Scale | Targets | PL | $K_p$ | DS | $\phi$ | $\theta$ |
|-------|---------|----|-------|----|-------|---------|
| 2     | Proposed| 5.04| 0.89  | 10.06| 4.29  | 9.84    |
|       | Bilinear| 31.69| 13.55 | 18.18| 23.75 | 27.57   |
| 4     | Proposed| 6.61| 1.19  | 13.61| 5.92  | 14.33   |
|       | Bilinear| 43.24| 18.75 | 24.94| 32.72 | 38.41   |
| 8     | Proposed| 8.30| 1.55  | 16.94| 8.00  | 18.13   |
|       | Bilinear| 51.15| 22.46 | 30.23| 39.27 | 46.2    |

TABLE VI: Cumulative Performance Gain of MAE and SDE Compared to Baseline Model

| MAE    | SDE    |
|--------|--------|
| + DA   | +136%  |
| + RES  | +26%   |
| + MTL  | 0      |

STL: Single task learning
RES: Residual connection and iterative up-and-down technique
DA: Data augmentation
A multi-task learning model for super resolution of wireless channel characteristics

Wang, Xiping

IEEE

Wang X, Zhang Z, He D, et al., (2023) A multi-task learning model for super resolution of wireless channel characteristics. In: 2022 IEEE Globecom Workshops (GC Wkshps), 4-8 December 2022, Rio de Janeiro, Brazil, pp. 952-957
https://doi.org/10.1109/GLOBECOM48099.2022.10001700

Downloaded from Cranfield Library Services E-Repository