Radio propagation prediction using deep neural network and building occupancy estimation

Kazuya Inoue\textsuperscript{1}, Koichi Ichige\textsuperscript{1, a)}, Tatsuya Nagao\textsuperscript{2}, and Takahiro Hayashi\textsuperscript{2}

\textsuperscript{1} Department of Electrical and Computer Engineering, Yokohama National University, Yokohama-shi, 240-8501, Japan.
\textsuperscript{2} KDDI Research Inc., Fujimino-shi, Saitama 356-8502, Japan.
\textsuperscript{a)} koichi@ynu.ac.jp

Abstract: In this paper, we propose a radio propagation prediction method using machine learning and building occupancy estimation. There have been learning-based researches using aerial photographs and building occupancy images as spatial information. However, the availability of building occupancy images are often limited to urban areas. We assume the situation that only the aerial photographs are available, and aim to improve radio propagation prediction accuracy by estimating building occupancy images from the given aerial photographs.

Keywords: radio propagation prediction, machine learning, neural network

Classification: Antennas and Propagation

References

[1] T. Imai, K. Kitao, and M. Inomata, “Radio propagation prediction model using convolution neural networks by deep learning,” Proc. European Conference on Antennas and Propagation, pp. 1–5, April 2019.

[2] T. Hayashi, T. Nagao, and S. Ito, “A study on the variety and size of input data for radio propagation prediction using a deep neural network,” European Conference on Antennas and Propagation, 2020. DOI: 10.23919/EuCAP48036.2020.9135876

[3] T. Nagao and T. Hayashi, “Study on radio propagation prediction by machine learning using urban structure maps,” European Conference on Antennas and Propagation, 2020. DOI: 10.23919/EuCAP48036.2020.9135353

[4] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: convolutional networks for biomedical image segmentation,” Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention, vol.9351, pp. 234–241, Oct. 2015. DOI: 10.1007/978-3-319-24574-4_28

[5] A. Krizhevsky, I. Sutskever, and G.E. Hinton, “ImageNet classification with deep convolutional neural networks,” Proc. Advances in Neural Information Processing Systems, pp. 1097–1105, 2012.
1 Introduction

Radio propagation prediction plays an important role in high-speed wireless communication. There are two representative but different methods in radio propagation prediction: experimental methods and theoretical methods. Recently, the prediction methods using machine learning have been developed [1, 2, 3]. The prediction model in [1] evaluates simulation data obtained from ray-tracing method, while the model in [2, 3] evaluates the data acquired in urban areas. The method in [2] employs the input data of the spatial data like map images, and the system parameters like base station specifications and Tx-Rx (transmitter-receiver) distance. It has been reported that using building occupancy images as spatial information can provide higher accuracy than using aerial photographs [2]. However, the building occupancy images are often provided only in urban areas.

In this paper, we propose a novel method to improve the accuracy when using only aerial photographs. We generate building occupancy images from aerial photographs using the Convolutional Neural Network (CNN) model called U-Net [4]. In addition, we also study the effect of using the midpoint image between Tx and Rx as additional spatial data.

2 Radio propagation environment prediction by neural network

Two of the authors have already developed a receiver power prediction method based on CNN and fully connected neural network (FNN) [2]. The spatial data and system parameters were used as the input data in [2], where aerial photographs and building occupancy images are used as spatial data. The distances between Tx and Rx, base station specifications and some other data are used as system parameters.

They employed the CNN based on AlexNet [5] whose model has four convolutional and two pooling layers. The FNN consists of input, output and six hidden layers, where the number of the weight parameter $W_\ell$ in the $\ell$-th layer is defined by

$$W_\ell = 4096 \times \left(\frac{1}{4}\right)^{\ell-1}.$$  \hspace{1cm} (1)

The final output of the prediction model is given as the received power in dBm, and is evaluated by the Root Mean Squared Error (RMSE) between the measured and predicted values.

3 Proposed method

We propose a radio propagation prediction method basically based on [2] but to improve prediction accuracy when using only aerial photographs, i.e., the building occupancy images are not available. We try to estimate building occupancy images from aerial photographs and use them in radio propagation prediction.

3.1 Occupancy image generation

It is indicated in [2] that the prediction accuracy becomes higher when using the building occupancy images as a spatial data, than using only aerial photographs. The Building occupancy images can make clear boundaries of buildings and roads. However, the areas where the building occupancy images are provided is limited...
only to urban areas and it is often difficult to get the latest image data. Therefore, we try to estimate building occupancy images from aerial photographs using the CNN and use them in radio propagation prediction. We adopt U-Net [4] for building occupancy image generation.

3.2 Radio propagation prediction

We performed radio propagation prediction using spatial data and system parameters as input data. As the spatial data, we use the rectangular images at three points, where the center of the images are given by the Tx point, the Rx point and the midpoint of Tx and Rx. The midpoint image compensates the spatial data between Tx and Rx. There are three types of images: aerial photographs, building occupancy images, and estimated building occupancy images generated by U-Net. As system parameters, we use 9 kinds, 12 parameters such as Tx antenna height, Tx antenna direction and directivity, Tx antenna gain, transmission power and Tx-Rx distance.

The prediction model consists of CNN and FNN. The CNN model has the convolution and pooling layers and they extract feature values. Its structure is same as that of [1]. In the FNN, the feature values of images and system parameters are inputted, and the final output is the predicted value of the received power in dBm. In network learning, MSE (Mean Squared Error) was used for the loss function and Adam was used for the optimization algorithm.

4 Simulation

We evaluated the proposed method using the measured LTE signals in 2.1GHz band obtained in Tokyo metropolitan area. The height of the transmission points are within 13 to 115m high, installed on the top of buildings. We obtained the measurement data by a receiver antenna installed on a roof of a running vehicle.

4.1 Dataset

After the data is measured by the above method, the median value within $5 \times 5$ m$^2$ area is extracted to remove instantaneous value fluctuation. When the received data contain the signals from multiple Tx points, the signal with the highest receiving power is taken as the data at that point. As a result, the total number of measurement data was 41,650 points. The aerial photographs and building occupancy images are created based on the location of the measured data. The image size is 64 x 64 pixels and covers an area of 256m$^2$. We use 4,000 points data for building occupancy image generation, and the remaining 37,650 points for radio propagation prediction. For both the building occupancy image generation and the radio propagation prediction, we use 90% of the total data as learning data, and the rest for test.

We prepared two types of data sets: Dataset 1 and Dataset 2. The test data in Dataset 1 is randomly extracted from all the measurement data, means that both the training and test data covers the whole area. Dataset 1 is constructed in the same way as in [2], but the generalization performance of the predictive model cannot be properly evaluated because of the mixture of training and testing data in the measurement area. Therefore, Dataset 2 is prepared in addition. Dataset 2 separates
the test and training data areas, i.e., those areas do not overlap to each other. It is obvious that the Dataset 2 is more difficult situation.

4.2 Occupancy image generation

We used the U-Net model described in Section 3.1 to train the building occupancy image generation model. The input image size is $64 \times 64 \times 3$ and the output image size is $64 \times 64 \times 1$. The network structure is similar to that of [4], with the upsampling and downsampling four times each. In addition, the number of channels in the input/output section is changed so that the input is an RGB image and the output is a grayscale image. Also the padding is done so that the image sizes of the input/output become same. The test data (400) is 10% of the total 4,000 points while the rest is used as the training data (32,400). In training the model, the mini-batch size is 32, the maximum number of epochs is 150, and the early stopping is applied when no improvement is seen in 50 epochs. After learning the model, the building occupancy images were generated from the aerial photographs using for radio propagation prediction.

Figure 1 shows the examples of (a) aerial photograph, (b) building occupancy image, and (c) estimated building occupancy image generated from the aerial photograph by U-Net. From Fig. 1, we see that the estimated image in Fig. 1(c) has a similar feature with the real building occupancy image in Fig. 1(b). Note that the building occupancy image in Fig. 1(b) is originally a binary image in which the place where the building is located is white while the other areas are black. The estimated image in Fig. 1(c) is a grayscale image that represents the probability of the building existence in each pixel. The colored aerial photograph in Fig. 1(a) is converted to grayscale when it is used for building occupancy image generation and radio propagation prediction.

4.3 Radio propagation prediction

We evaluated the proposed method by performing radio propagation prediction using Dataset 1 and Dataset 2 shown in subsection 4.1. The evaluation index is the RMSE between the measured and the predicted receiver powers, and is expressed calculated by

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (p_i - \tilde{p}_i)^2}$$

(2)
where $N$ is the number of data, $p_i$ is the measured receiver power value, and $\tilde{p}_i$ is the predicted receiver power value. In training the model, the mini-batch size is 648, the maximum number of epochs is 1000, and the early stopping is applied when no improvement is seen in 50 epochs. We also compared the RMSE values of the input image with and without the midpoint image to evaluate the effect of adding the midpoint image. The results for each type of input data is shown in Table I, where the methods 1, 2 and the proposed method used aerial photographs, true building occupancy images, and the estimated building occupancy images, respectively. We confirmed from Table I that the RMSE value became smaller when the spatial information (images) was added to the system parameters than when only system parameters were used for prediction. Besides in Dataset 2, the RMSE value was improved by using the estimated building occupancy image and adding the midpoint

| Dataset  | Input type                     | Method 1 | Method 2 | Proposed Method |
|---------|-------------------------------|----------|----------|-----------------|
| Dataset 1 | Rx, Tx images + System Parameter | 4.10     | 3.93     | 4.01            |
|         | Rx, Tx, Mid images + System Parameter | 4.06     | 3.93     | 3.91            |
|         | System Parameter Only         |          |          | 4.49            |
| Dataset 2 | Rx, Tx images + System Parameter | 8.99     | 8.19     | 8.36            |
|         | Rx, Tx, Mid images + System Parameter | 7.81     | 7.66     | 7.54            |
|         | System Parameter Only         |          |          | 9.06            |

![Example of received power prediction result](image_url)

**Table I.** Comparison of RMSE (in dB)

**Fig. 2.** Example of received power prediction result
image, compared to the case where the original aerial photographs were used as spatial information. This is because we could capture the positions of buildings from the estimated building occupancy images.

As an example of the radio propagation prediction results, Figs. 2(b), 2(c), and 2(d) show the difference between the estimated and measured received powers on a map of Dataset 2, while Fig. 2(a) shows the measurement data as a reference. Here, the error is defined as the difference of the measured value from the predicted value. We see from Figs. 2(b) and 2(c) that there are many red and blue points (with large estimation error) but many white points (with small estimation error) in Fig. 2(d). That means, the use of the midpoint image is effective in the received power prediction.

5 Concluding remarks

We proposed a method of estimating building occupancy images from aerial photographs using U-Net for accurate radio propagation prediction. The RMSE of the received power estimation became better when using the images generated by U-Net than the case of aerial photographs only.