IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 5, 2021, accepted March 26, 2021, date of publication April 2, 2021, date of current version April 30, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3070711

Millimeter Wave Path Loss Modeling for 5G Communications Using Deep Learning With Dilated Convolution and Attention

HONG CHENG[1], SHENGJIE MA[1], HYUKJOON LEE[1], AND MINSUNG CHO[2]
[1]Department of Computer Engineering, Kwangwoon University, Seoul 01897, South Korea
[2]Network Division, Network Automation Group, Samsung Electronics Company Ltd., Suwon 16677, South Korea

Corresponding author: Hyukjoon Lee (hlee@kw.ac.kr)

This work was supported in part by the Ministry of Science and Information and Communication Technology (ICT) under Grant NRF-2016R1D1A1B04936096, and in part by the Research Grant of Kwangwoon University, in 2020.

ABSTRACT An accurate and efficient path loss modeling method for millimeter wave communications plays a significant role in the large-scale deployment of a fifth-generation (5G) mobile communication system. Conventional path loss modeling methods such as deterministic methods, empirical methods, and machine learning-based methods, in practice, cannot achieve the desired level of performance in terms of accuracy. This paper proposes a novel path loss model for 5G communications in suburban scenarios using deep learning with dilated convolution and attention. Dilated convolution is used to alleviate the locality of feature extraction and capture the global information on input images. Attention rendered by global context blocks helps the attention-enhanced convolutional neural network (AE-CNN) model utilize the global information on inputs to extract essential features of propagation environments. A distance-embedded local area multi-scanning algorithm which generates input images that can improve learning the latent features with dependency on distance is proposed. The experimental results indicate that the AE-CNN model can outperform state-of-the-art deterministic and empirical methods in terms of root mean square error in test scenarios.

INDEX TERMS Deep learning, 5G mmWave, path loss modeling, dilated convolution, attention.

I. INTRODUCTION

Millimeter wave (mmWave) communication is one of the top-priority candidate radio technologies for the 5G mobile systems [1]. Due to large bandwidth and high carrier frequency, mmWave communications are characterized by smaller scattering effect and more significant blocking effect of non-line-of-sight (NLoS) paths [2]. Path loss modeling of mmWave communications is challenging due to increased sensitivity to propagation environments. Therefore, path loss modeling for mmWave propagation plays a vital role in designing and analyzing 5G communication systems. Three types of conventional path loss modeling methods have been investigated in previous studies, namely empirical methods [3]–[7], deterministic methods [8], [9], and machine learning-based (ML-based) methods [10]–[18].

The associate editor coordinating the review of this manuscript and approving it for publication was Giovanni Angiulli.

Empirical methods predict the path loss by calculating the parameters of fixed equations based on the statistical characteristics of data, such as carrier frequency and TR-separation, namely the distance between a transmitter (Tx) and a receiver (Rx). These equations are variations of the Friis free-space path loss model, and a few parameters are required [19]. Empirical methods such as the close-in free space reference distance (CI) model and alpha-beta-gamma (ABG) model [4] are relatively easy to handle when the measurement data are given. However, the parameters of empirical models should be determined based on the measurement data collected in a specific scenario and thus their prediction accuracy may decrease when they are applied to a scenario that differs considerably from the measured scenario.

Deterministic methods, such as the three-dimensional (3D) ray-tracing method [20], are based on physical optics simulation. They require detailed site-specific information, such as the configuration of a Tx and an Rx and the dielectric properties of the materials of building structures. This lowers

the prediction accuracy of the 3D ray-tracing methods in a large-scale deployment scenario since such information is rarely available in practice. Moreover, it suffers from a tremendous degree of computational complexity.

The path loss modeling can also be handled by data-driven approaches such as the ML-based methods as a regression problem, where the salient features calculated based on the locations of measurement and propagation environment data are used as inputs, and measured path loss values are used as desired outputs in training. ML-based path loss models trained for a small number of representative propagation environments can generalize to predict the path loss values in previously unseen sites. They can usually achieve better performance than empirical models, thanks to their ability to learn the underlying relationship between inputs and outputs. However, the input preprocessing, including feature selection and extraction, which requires domain expertise and an enormous amount of computation, has significant influences on the performance of prediction.

The unsatisfactory performance of conventional path loss modeling methods and their low efficiency in practice have prompted researchers to find an alternative path loss modeling method such as deep learning methods [21]–[23]. Deep learning is a subset of machine learning algorithms that use multiple processing layers to learn representations of data with multiple levels of autonomous feature abstraction [24].

In [23], for instance, it is shown that CNNs can be trained to predict path loss values based on the input images that are expected to encode the structural information of a local map for each pair of Tx and Rx locations. However, it is not clear whether the CNNs are trained to model the underlying path loss functions based on all the latent features which may not even be extracted as expected. Although the global features, such as the distance between a Tx and an Rx, LoS/NLoS, and the existence of large buildings and streets, are known to have a large impact on the path loss, a close inspection on the predicted values of the CNNs reveals they may have not been extracted or used successfully. This may be caused by the locality of feature extraction by the CNNs. Specifically, the locality is caused by the kernels in the convolutional layers of CNNs as the receptive field of a neuron in the kernels cannot cover the entire input image. The utilization of global information in path loss modeling means that the CNN model can extract or select features based on the overall information on the inputs or the information on an area as large as possible. Attention mechanism can increase the expressive ability of extracted features for CNN by emphasizing important parts and suppressing unnecessary features [25]–[30]. Dilated convolution supports an exponential expansion of the receptive field without loss of resolution or coverage [31].

This paper proposes a novel deep learning path loss model called the attention-enhanced convolutional neural network (AE-CNN) model for 28 GHz mmWaves in suburban scenarios. Dilated convolution is used to alleviate the locality of feature extraction and capture the global information on input images. Global context (GC) blocks [32] are used to utilize the global information on inputs to extract essential features of propagation environments. A distance-embedded local area multi-scanning (DE-LAMS) algorithm is proposed to generate input images to improve learning the latent features with dependency on distance. The proposed model achieves a satisfactory level of performance on path loss predictions in test scenarios. The proposed AE-CNN path loss model outperforms conventional empirical and deterministic methods, as demonstrated by experimental results.

The rest of this paper is organized as follows. Section II presents related works on ML-based path loss modeling methods, CNN-based path loss models, attention, and dilated convolution. Details of the proposed AE-CNN path loss model are described in Section III, including the illustration of input preprocessing, the DE-LAMS algorithm, and the AE-CNN model description. The experimental results are presented and discussed in Section VII. Finally, Section V presents the main conclusions drawn from this study.

II. RELATED WORKS

A. ML-BASED PATH LOSS MODELING METHODS

ML-based methods can learn features from statistical information of measurement data (as well as empirical models) or from preprocessed inputs based on knowledge of the radio propagation field to predict the path loss values. There is a tremendous amount of research on path loss modeling with ML-based methods, including neural networks, support vector regression, and decision tree-based methods [10]–[18]. The ability to model complex relationships between inputs and outputs makes ML-based methods suitable for path loss modeling.

Yang and Lee present a neural network-based path loss model with a series of feature extraction algorithms [14]. Their algorithms can extract features from digital elevation map (DEM) data as inputs of neural network models. ML-based methods can achieve good performance using the useful features extracted based on radio propagation knowledge in specific scenarios. However, 3D map data of studied areas such as the DEM data is required to extract useful features but rarely available in practice. Extensive computation of feature extraction is inevitable, and high computational resources are usually essential for input preprocessing of ML-based path loss modeling methods. These limitations motivate researchers to find better methods to extract features for ML-based path loss modeling methods. Hybrid-empirical path loss models that use ML methods to minimize the errors between the predicted path loss using empirical methods and measured path loss have been proposed in [10]–[12]. The ML methods in these hybrid-empirical models can be used to select or predict the parameters of empirical models. The inputs of these models usually include the inputs of conventional path loss modeling methods, such as TR-separation, the transmission power of Tx, and carrier frequency. The performance of hybrid-empirical models is usually better

than that of an empirical model but similar or slightly worse than that of ML-based methods because feature extraction is usually not used in hybrid-empirical models. Ensuring that the inputs provided contain enough appropriate information on the propagation environment for learning plays a decisive role in the performance of ML-based path loss modeling methods.

B. CNN-BASED PATH LOSS MODEL

In [21], a method of assigning images to path loss measurements is proposed to generate input images for a CNN model to estimate path loss in urban areas. The CNN model used has a simple structure with two convolutional layers and one hidden fully connected (FC) layer. The images are generated based on the height of buildings in the studied area and the location of Tx and Rx. Each image is captured from a rectangle area between Tx and Rx is resized to a square of predefined dimensions. However, the resize of images can destroy the original information on propagation environments, especially the distance information. The performance of CNN's prediction can be influenced as well. In [23], a CNN-based path loss model is proposed for mmWave path loss modeling in suburban environments. The CNN-based path loss model has four subnetworks to predict the path loss values of the four directional antennas of an Rx. Feature-sharing layers are used to share common knowledge between subnetworks and an additional distance neuron is used to provide explicit distance information. The enhanced local area multi-scanning (E-LAMS) algorithm extracts local area information on the path loss environment between each pair of Tx and Rx locations from the processed Google map images where buildings and streets are captured. The CNN model achieves better performance than state-of-the-art empirical models. However, close inspection shows that the prediction accuracy of the CNN-based model is not stable even in the case of LoS sites. We describe it in more detail in section III.

C. ATTENTION AND DILATED CONVOLUTION

Attention mechanisms can be adopted as an integral part of CNNs to capture global dependencies [28], [33]–[35]. Reference [34] describes the squeeze-and-excitation (SE) block, an architectural unit that can be plugged into CNNs to improve the performance of CNN models with only a slightly increased total number of parameters. SE blocks explicitly model channel relationships and channel interdependencies and include a form of self-attention on channels. In [35], the potential of using multi-layered attention is demonstrated. An attention module called convolutional block attention module (CBAM) with spatial attention and channel attention in image captioning tasks is proposed in [36]. The use of CBAM highlights the broad applicability of the attention, particularly for image classification and object detection tasks.

A lightweight attention mechanism provided by GC blocks is proposed in [32]. A GC block can calculate the positional relationship between the query position and other positions in input images. The GC block is designed to use a query-independent attention map for all query positions with a small computation cost. Fig. 1 shows the structure of the GC block. Input features extracted by convolutional layers can be reshaped as 3D or 2D tensors. The input feature with the shape of $C \times H \times W$ is used three times in the GC block. First, the input features are converted as a $1 \times H \times W$ tensor using 1-by-1 convolution. Then, it can be reshaped as a $HW \times 1 \times 1$ tensor. The input features are also reshaped as a tensor with the shape of $C \times HW$; then, this tensor is used to calculate the dot product using the Softmax of the $HW \times 1 \times 1$ tensor. The output of the dot product is a $C \times 1 \times 1$ tensor. The first two steps can be called context modeling that determines global information on the input features. The SoftMax of the reshaped feature is performed to obtain attention. Second, feature transformation is performed by a bottleneck transform module that includes two 1-by-1 convolution procedures to capture channel-wise dependencies. Third, broadcast element-wise addition is applied to aggregate the extracted global context features, i.e., the $C \times 1 \times 1$ tensor, to each neuron of the original input features.

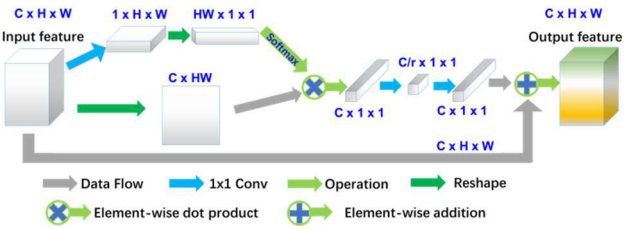


FIGURE 1. The structure of the GC block. C is the number of channels of the input feature. H and W are the height and width of the input feature, respectively. HW is a value of dimension that equals the product of H and W. r is the reduction ratio, and C/r denotes the dimension of a bottleneck.

Dilated convolution has been proposed to improve the performance of semantic segmentation as it supports an exponential expansion of the receptive field without loss of resolution or coverage [31]. Fig. 2 shows the general convolution and dilated convolution. Fig. 2 (a) indicates a 5×5 input is convolved with a 3×3 kernel with stride $s = 2$ and padding $p = 1$ to produce an output of 3×3 by the general convolution. Fig. 2 (b) shows a 7×7 input is convolved with a 3×3 kernel with $s = 1$ and $p = 0$ to produce an output of 3×3 by the dilated convolution. Fig. 2 indicates that dilated convolution can increase the size of the receptive field with the same kernel size, namely, the same number of trainable parameters.

With GC blocks and dilated convolutional layers, CNNs can utilize global information on the extracted features or inputs and focus on useful features at the cost of small increase in the number of parameters for calculation of attention.

III. AE-CNN PATH LOSS MODEL

Close inspection in the contribution of input values to the predicted path loss values of the CNN model obtained in [23]

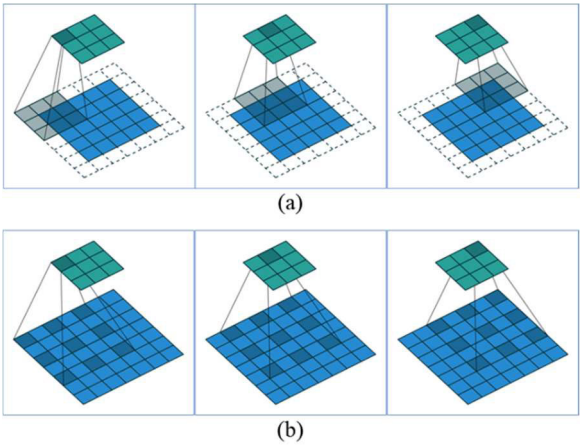


FIGURE 2. General convolution (a) and dilated convolution (b). Blue squares are inputs, and cyan squares are outputs. The interconnected neurons in each step are marked in a darker color.

has suggested CNNs may not have been completely successful in approximating the underlying path loss model. The first observation made is that it is difficult to find a correlation between the path loss predictions and TR-separation, as shown in Fig. 3 (b), where the path loss values are not only similar regardless of the TR-separation values but the predictions of two legs are asymmetrical as well. This implies the CNN-based model trained with the TR-separation values fails to learn the dependency of path loss values on them.

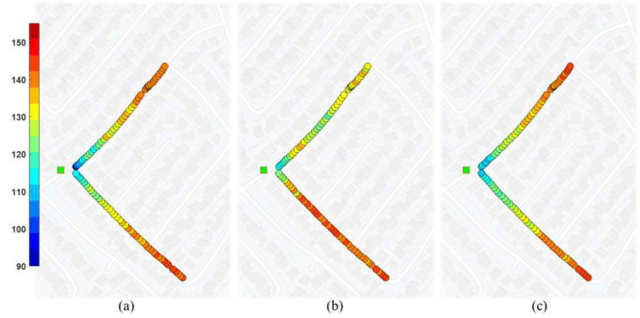


FIGURE 3. Measurement data (a), predicted values of the CNN-based model trained with E-LAMS images and distance neurons (b), and predicted values of the proposed AE-CNN model trained with DE-LAMS images (c) in S7. The green point indicates the Tx location, and the points with distinct colors show the path loss values in the measured locations of Rx, as shown in the color bar on the left.

The second issue is that the parameter values of the trained CNNs do not conform very well to the basic principles of mmWave path loss models as visualized by the heatmap values of extracted features (Fig. 4). In what follows, we address these two issues in more detail and propose new approaches to CNN-based path loss modeling of mmWave.

A. DE-LAMS ALGORITHM FOR INPUT IMAGE GENERATION

In theory, the amount of path loss values of the signal from a Tx to an Rx is given by $P = 10n \log_{10} d + C$, where P is the path loss in dB, n is the path loss exponent (PLE), d is TR-separation in meters, and C is a variable that accounts for system losses. This implies the parameters associated with the distance neuron are not updated thoroughly enough to influence the path loss prediction significantly.

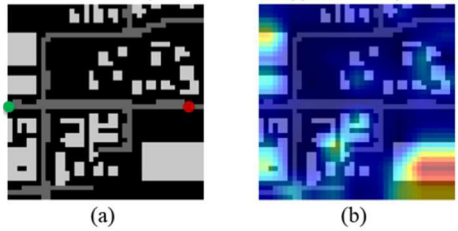


FIGURE 4. An E-LAMS image (a) and its overlapped heatmap (b) of a CNN model. The green and red points show the locations of Tx and Rx in (a). The color of the heatmap overlapped on the buildings and streets in the E-LAMS image indicates the importance of the extracted features by the CNN model. Specifically, the red color of the heatmap indicates that the covered area contains crucial features, whereas the blue color indicates that the covered area contains unimportant features.

Algorithm 1 DE-LAMS Algorithm

Input:
GPS locations of Tx and Rx
Size of a DE-LAMS image: (w, l)
Processed Google map image matrix M_G

Output:
A DE-LAMS image

1. Calculate the value of background pixels, p_M
2. Determine an extended local area A_E on M_G
3. **for** $i = 1$ to l **do**
4. Determine the equation and location of the i_{th} scan line on A_E
5. Select w pixels on the i_{th} scan line
6. Copy w selected pixels from M_G as an array s_i
7. **end for**
8. Integrate l arrays from s_1 to s_l as a matrix M_I
9. Set values of background pixels in M_I as p_M
10. Save M_I as a DE-LAMS image

In order to address this issue, we propose to include the distance information in the training input images for a CNN to take the distance into account more efficiently. As presented in Algorithm 1, the distance-embedded local area multi-scanning (DE-LAMS) algorithm generates CNN input images with the distance information embedded. The local area is a squared area that takes the straight line between Tx and Rx as the central axis with an extension of 10% margin beyond the Rx. The extended area is included to allow for the influence of objects around the Rx with respect to the reflection in propagation environments. The main buildings and streets are kept in the input images. Each image generated by this algorithm can be regarded as an encoded image with respect to a pair of Tx and Rx. Multiple scan lines parallel to the Tx–Rx line are determined. The scan lines should be able

to slice the local area into multiple pieces having the same length and width. The values of pixels on each scan line at a predefined interval can be selected and copied to form a vector. These vectors of all scan lines can be integrated to form a matrix. Before saving the matrix as an input image, the pixel values of the background pixels, i.e., pixels other than those representing buildings and streets, are changed according to the values of TR-separation. The input images are called DE-LAMS images, where the base pixel values change proportionally to the values of the TR-separation in the local area.

B. THE ARCHITECTURE OF AE-CNN MODEL

Feature extraction of CNNs is performed by the kernels of convolutional layers. Since the receptive field of a neuron in convolutional layers is always smaller than the size of the input images. This makes CNNs learn local information on input images and fail to utilize the global information (i.e., all information exists in the input images).

Each building and street in a propagation environment should have a different amount of impact on the path loss. For instance, the buildings close to Tx and Rx should have a larger influence on the path loss than those in the middle of Tx to Rx. If we consider an extreme circumstance where there is only a big building lies in the path from Tx to Rx and the signal is blocked, the CNN can end up arriving at a local minimum with high activation values for the parameters of this building, although the influence of this building should be small. This results from the loss of useful global information owing to the locality of feature extraction. This can be manifested by the heatmap of a CNN-based model presented along with the E-LAMS image (Fig. 4), the buildings along the left edge, at the lower right corner, and in the center receive more attention than the streets. It can also be noticed that some areas with no buildings and streets, such as the lower right corner of Fig. 4 (b), are covered with yellow color, while most streets are covered with blue color.

We further ascribe the improper utilization of global information to the locality of feature extraction. The key to achieving better performance is to extract useful features that contain more global information on input images since path loss modeling for mmWave is sensitive to the location information of buildings and streets on path loss environments. The discovery of buildings that have significant influences on path loss is also important.

GC blocks enable channel attention based on the spatial representation to be learned from the global context of the input features for better feature extraction. Features such as buildings and streets that significantly influence the path loss can be emphasized, whereas unimportant features are suppressed. Dilated convolution can help a network capture more information on inputs with a minimal number of additional parameters to increase the size of the receptive field of neurons. Therefore, we propose to use GC blocks and dilated convolutional layers that can be augmented in a conventional CNN model.

TABLE 1. Parameters of dilated convolution layers in The AE-CNN path loss model.

Number of convolutional layer	1st	2nd	3rd	4th
Input size	41	19	9	7
Input channel	3	8	6	3
Output size	19	9	7	3
Output channel	8	6	3	4
Kernel size	5	3	3	3
Stride size	2	2	1	2
r of GC block	4	2	3	2

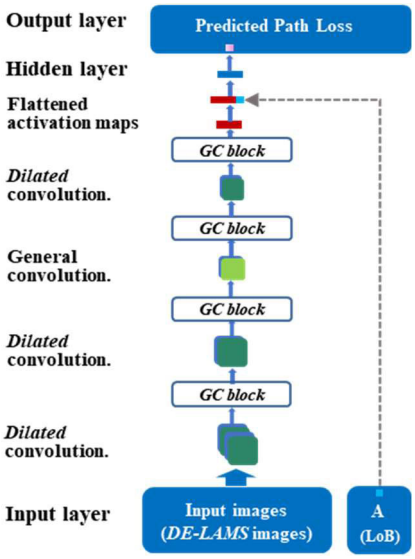


FIGURE 5. Network structure of the proposed AE-CNN path loss model.

Fig. 5 shows the network structure of the proposed AE-CNN model. The feature extraction part of the AE-CNN model is composed of four dilated convolutional layers and three GC blocks. We interleave GC blocks between convolutional layers. The parameters of convolutional layers in the model are presented in Table 1. When the stride size of one convolutional layer is one, we cannot use dilated convolution in that layer. Thus, the third convolutional layer uses general convolution. The FC layers, including one hidden layer of 20 neurons, predict path loss values using the features extracted from DE-LAMS images with an angle neuron A in the input layer. The neuron A contains the values of the line of bearing (LoB) [37]. The LoB of an Rx is the absolute angle between the direction from Tx to Rx. The LoB indicates the propagational path and direction of mmWave signals and can affect path loss when directional antennas are used in the Tx and Rx. The outputs are the predicted path loss values.

The main idea of this work is to capture as much global information and essential information from inputs as possible. The GC block learns a spatial representation of the global context from the input features to obtain better channel attention for helping the feature extraction of upper layers. Dilated convolution can alleviate the locality of feature extraction by

capturing features in a larger receptive field than the general convolution.

IV. EXPERIMENTS

We first describe how the input images are generated based on the measurement data collected and the map images of the regions where we collected data.

A. PREPROCESSING

1) MEASUREMENT DATA COLLECTION AND PREPROCESSING

Field measurements were carried out in two areas of the United States by Samsung Electronics. A total of 17 scenarios (S1–S17) were included in the field measurements. Measurements of 13 scenarios (S1–S13) were conducted in California, and the remaining four scenarios (S14–S17) were conducted in Houston. Moreover, the simulated path loss values using a 3D ray-tracing method in S14–S17 were provided with the measured path loss values. In each measurement scenario, a 28 GHz Tx with a directional antenna and a known location was fixed on the ground, and a 28 GHz Rx was set on the rooftop of a moving car. For a Tx with a known location in a scenario, measurement data including the locations of the Rx and the receiver signal strength (RSS) values or path loss values of a directional antenna(s) of the Rx were recorded at a certain time interval when the car was driven on designated routes.

The Rx used in S1–S13 has four directional antennas, whereas that used in S14–S17 has just one directional antenna. We calculated the average RSS values of four antennas as the RSS value in each Rx measurement location in S1–S13. The averaged RSS value was then converted to the path loss value by subtracting the RSS value from the transmission power of the Tx in that scenario. The transmission power of the Tx in S1–S7 and S8–S13 was 39 dBm and 53 dBm, respectively.

2) MAP IMAGE PROCESSING

CNN input images containing information on propagation environments should be generated to train a CNN based on the measurement data and the map data. Two-dimensional (2D) Google street map tiles in the areas where we conducted the field experiment were downloaded. These tiles were stitched into complete map images for each measurement scenario. The buildings and the objects around streets in suburban scenarios are assumed to be the pivotal factors in the propagation environments that have significant influences on path loss. A series of image processing methods were used to extract buildings and streets from the 2D Google streets map images. First, the map images were converted into grey images. Then, the buildings and streets were extracted by selecting these pixels based on their pixel values. The Gaussian filter-based image smoothing method was applied to remove noisy pixels. Finally, we obtained the processed Google map images with extracted buildings and streets.

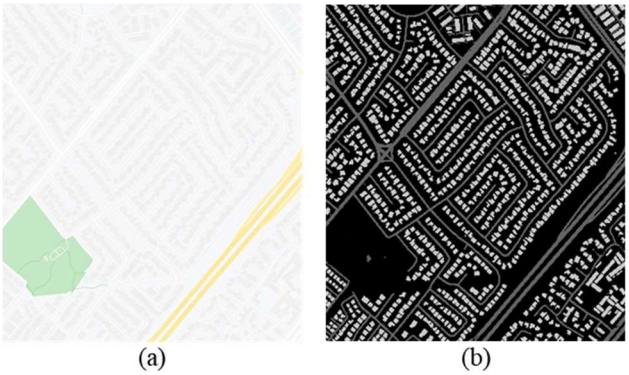


FIGURE 6. An example of a processed map in S7. (a) Google map image. (b) Processed Google map image. The pixel values of buildings and streets are set as 200 and 100, respectively, for better visualization.

Fig. 6 shows a Google street map image and a processed Google map image. The grey-level pixel values of buildings and streets used for the input image generation are set as 20 and 10, respectively.

3) HMM-BASED MAP MATCHING

There are some Rx locations of measurement data aligned off roads due to measurement errors caused by the limitation of GPS equipment and the positioning method, as shown in Fig. 7 (a). This makes the Rx locations of measurement data imprecise. The alignment of GPS locations with the roads is called map-matching [38]. The road is divided into two lanes since a car is driven back and forth on designated routes. In this work, we implemented a lane-level hidden Markov model (HMM)-based map-matching method [39] to align the outliers of the Rx locations on the roads. The Rx location offsets caused by the GPS error conform to a Gaussian distribution [40], and most of the offsets are concentrated in the Gaussian mean. In this work, we first applied the lane-level HMM-based map matching to calculate the adjustment vector of each Rx location. Then, the average adjustment vector of all adjustment vectors was added for all Rx locations of measurement data.

The distribution of GPS error and the position relationship between each Rx location can be maintained by moving all Rx locations of measurement data together onto the road,

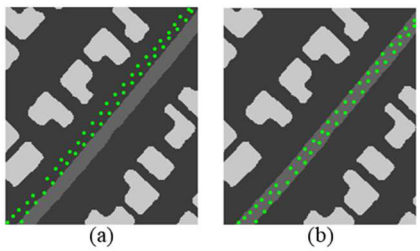


FIGURE 7. An example of the lane-level HMM-based map matching. (a) Actual Rx locations of measurement data. (b) Corrected Rx locations of measurement data. The Rx locations of the measurement data are plotted as green points on the preprocessed Google map image.

as shown in Fig. 7 (b). The Rx locations of measurement data used in each scenario were corrected by the average adjustment vector of the lane-level HMM map-matching method. The measured path loss values in S17 are visualized by overlapping Rx locations on the Google map image. As shown in Fig. 8, S17 is a scenario where all Rx points are LoS sites.

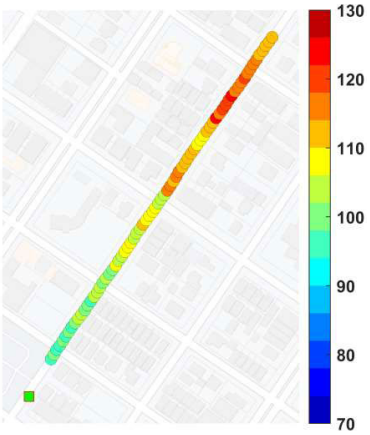


FIGURE 8. Visualization of measured path loss values in S17. The green square with a red edge indicates the location of Tx. The points on the road are the locations of Rx, and the path loss values in dB are indicated by the filled color as shown in the color bar.

4) INPUT IMAGE GENERATION

E-LAMS images are used as the main components of the inputs of the basic CNN model. We show examples of E-LAMS images in Fig. 9. The TR-separation and LoB are calculated based on the GPS locations of Tx and Rx after map-matching. The range of TR-separation values used in the experiments is [30, 600].

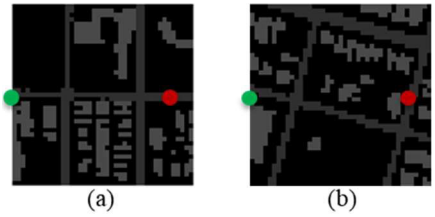


FIGURE 9. Examples of E-LAMS images, where the green and red dots show the locations of Tx and Rx, respectively.

With the calculated TR-separation and DE-LAMS algorithm, DE-LAMS images are generated as examples shown in Fig. 10, where the TR-separation of four DE-LAMS images increases from (a) to (d). The pixel values of buildings and streets are fixed at 120 and 100. The range of pixel values of the background pixels is [1, 80].

Upon completion of the image generation, the proposed AE-CNN path loss model was trained and its performance was evaluated. Details of the experimental settings and dataset separation are presented in subsection IV-B. Three sets of experiments were conducted as presented in

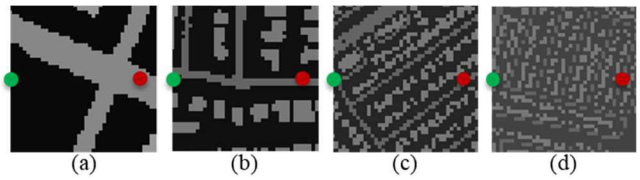


FIGURE 10. Examples of DE-LAMS images, where the green and red dots show the locations of Tx and Rx, respectively. The grey level of the background pixel indicates the values of normalized TR-separation.

subsection IV-C. First, experiments on the usage of distance information in inputs were conducted. Second, experiments were conducted to show the performances of the combinations of functional modules in the proposed AE-CNN model. Third, experimental results of conventional deterministic and empirical methods are presented. Comprehensive visualization with detailed discussion and analysis is presented in subsection IV-D. First, we visualized the extracted features to show the influence of DE-LAMS images on feature extraction. Second, the predicted values of AE-CNN models is compared with measurement data and the predicted values of conventional methods to show the performance of the proposed model. Third, the predicted errors of AE-CNN on a test scenario are visualized.

B. EXPERIMENTAL SETUP

The experiments were conducted using Keras and TensorFlow [41]. The stochastic gradient descent (SGD) optimizer was used for the experiments. The detailed hyperparameters of the models are listed in Table 2, and statistical information on the dataset separation is presented in Table 3. There are 8502 samples in the training dataset and 1237 samples in the test dataset, including S7, S13, and S17.

TABLE 2. Settings of hyperparameters.

Hyperparameters	Settings
Learning rate	0.001
Batch size	64
Maximum training epoch	1000

TABLE 3. Data separations.

Data	Scenarios	Number of data item
Training	11 scenarios in California 3 scenarios in Houston	8502
Test	S7	472
	S13	465
	S17	300

C. EXPERIMENTAL RESULTS

Experiments were conducted to show the performances of functional modules proposed in the AE-CNN model. The root mean square error (RMSE) of predicted path loss in test scenarios are presented in Table 4.

TABLE 4. Results of experiments.

Sets	Model	Inputs [c]	RMSE in test scenarios [d]		
			S7	S13	S17
A	CNN	E-LAMS	8.99	10.95	11.75
	CNN	E-LAMS + TR	8.39	9.86	11.48
	CNN	DE-LMAS	**6.44**	**8.97**	**9.42**
B [a]	D-CNN	DE-LAMS	5.29	7.94	4.52
	CNN+GC	DE-LAMS	5.22	7.63	5.24
	AE-CNN	DE-LAMS	***4.64***	***7.18***	***3.85***
C [b]	CI	-	**17.53**	**25.69**	29.33
	ABG	-	17.56	25.94	30.02
	3D ray-tracing	-	-	-	**3.96**

[a]The basic CNN model with dilated convolutional layers and GC blocks is denoted as "D-CNN" and "CNN+GC", respectively.

[b]The RMSE of the CI and ABG model on the training dataset of deep learning models is 15.99 and 16.06, respectively. The simulated results of the 3D ray-tracing method are denoted as "3D ray-tracing".

[c]"Inputs" includes "TR", "E-LAMS", and "DE-LAMS", denoting that the models are trained with the distance neuron, angle neuron, E-LAMS images, and DE-LAMS images, respectively. The values of LoB are used as the angle neuron A in the models of sets A and B. The LoB is not shown in the inputs for brevity.

[d]The bold values indicate that the model achieves the best performance in that experimental set for that scenario, whereas the values with italic and bold fonts indicate that that model achieves the best performance for that scenario.

1) USAGE OF DISTANCE INFORMATION

As described in subsection III-A, the DE-LAMS algorithm generates DE-LAMS images with the distance information embedded. We conducted experiments to validate that the DE-LAMS images can help CNN models achieve better performance than other inputs, as presented in set A of Table 4. The basic CNN models were trained with E-LAMS images, E-LAMS images and the distance neurons, and DE-LAMS images, respectively. It can be observed that CNN models trained with the DE-LAMS images achieved smaller RMSE than the other models. This indicates the CNN models trained by DE-LAMS images can efficiently learn the influence of latent features with dependency on distance. Closer inspection is presented in subsection IV-D with various visualizations.

2) DILATED CONVOLUTION AND GC BLOCKS

Dilated convolution can capture more global information from inputs to alleviate the locality of feature extraction. GC blocks can obtain the attention of prominent features of propagation environments from inputs. As presented in set B of Table 4, the performance of the D-CNN model, CNN+GC model, and the proposed AE-CNN model are studied. These models were trained with DE-LAMS images. As one can see, the D-CNN model performs better than the basic CNN model in terms of the test RMSE. This indicates that the global information extracted from input images using dilated convolution is beneficial to the performance improvement of the CNN models. The CNN+GC model outperforms the D-CNN models except in S17. As all Rx locations of measurement data in S17 are LoS sites, the D-CNN model can benefit from its larger receptive field of the neurons in dilated convolutional layers. The experimental results indicate that the D-CNN model can capture more useful environmental information on the streets than the CNN+GC model. The AE-CNN model can capture global information from inputs to alleviates the locality of feature extraction with dilated convolutional layers. The GC blocks of the AE-CNN model can help the AE-CNN model extract essential features having significant influences on path loss and suppress unnecessary features. Thus, the proposed AE-CNN model can achieve satisfactory performance for all test scenarios.

Ablation experiments were conducted to study the influences of different numbers of dilated convolutional layers and GC blocks on the performance of CNN models. The results are presented in Tables 5 and 6. Table 5 presents the RMSE of models with respect to the number of dilated convolutional layers. The number of dilated convolutional layers means how many times we replace a general convolutional layer with a dilated convolutional layer. There are three dilated convolutional layers in the AE-CNN model, and its third convolutional layer uses the general convolution, as shown in Fig. 5. Table 6 presents the RMSE of models with respect to the number of GC blocks in S7. The number of GC blocks means how many times we add a GC block to a CNN model or a D-CNN model after a convolutional layer. As presented in Tables 5 and 6, the RMSE is smaller when more dilated convolutional layers and GC blocks are used in CNN models. It indicates that the use of them together makes the AE-CNN path loss model achieve better performance than the models that use just one of them.

TABLE 5. RMSE of ablation experiments on number of dilated convolutional layers.

Models	Number of dilated convolutional layers		
	1	2	3
CNN	5.47	5.42	5.29
CNN+GC	5.17	4.82	4.64

TABLE 6. RMSE of ablation experiments on number of GC blocks.

Models	Number of GC blocks			
	1	2	3	4
CNN	5.74	5.52	5.34	5.22
D-CNN	5.20	5.10	4.79	4.64

3) RESULTS OF CONVENTIONAL DETERMINISTIC AND EMPIRICAL METHODS

Experimental results of conventional methods are presented in set C of Table 4, including the empirical CI model and ABG model and the deterministic 3D ray-tracing method. The parameters of the CI and ABG model calculated by solving the closed-form solutions via mathematical derivations are presented in Table 7. Samsung Electronics provided the simulated path loss values using a 3D ray-tracing method

TABLE 7. Parameters of empirical methods.

Models	CI model		ABG model	
Parameters[a]	n	3.04	α	2.57
			β	23.85
			γ	3.36
	σ	16.06	σ	15.99

[a] n is the PLE. α, β, and γ is a frequency-dependent term, a floating-offset value, and a distance-dependent term, respectively. σ is the standard deviation of the fading term.

with the measurement data. The experimental results indicate that CNN models outperform conventional methods in terms of the RMSE in the test scenarios. The proposed AE-CNN path loss model even achieves slightly better performance than the state-of-the-art deterministic 3D ray-tracing method in S17.

D. VISUALIZATION AND DISCUSSION

Various visualizations are presented to validate the satisfactory performance of the proposed AE-CNN model by solving the issues of the basic CNN-based path loss model for 5G communications in suburban environments.

1) VISUALIZATION OF EXTRACTED FEATURES

Features of propagation environments captured from input images by the proposed AE-CNN model are visualized as shown in Fig. 11. The input image and its overlapped heatmap are shown together to demonstrate the effects of the DE-LAMS images and the attention rendered by GC blocks. As shown in the second column of Fig. 11 (a), the bottom right building is considered as the main feature of propagation environments. This indicates that the CNN model cannot extract useful information sufficiently enough from E-LAMS images. Moreover, the streets are not thought of as crucial factors in the propagation environment. The images in the third column of Fig. 11 (a) show that the GC blocks emphasize the importance of buildings in the bottom center and buildings on the top side. E-LAMS images can just have a small influence on the feature extraction of the AE-CNN model. When the inputs, namely the E-LAMS images, can only provide a limited amount of information on propagation environments, the attention cannot exhibit a significant effect in the feature extraction.

Fig. 11 (b) shows the overlapped heatmaps of the AE-CNN model trained with DE-LAMS images. The images in the second column of Fig. 11 (b) show more buildings are overlapped with red color compared with Fig. 11 (a), and some streets are overlapped with light blue color as marked by the purple boxes. It indicates that more buildings and streets are considered to influence path loss. The importance of the buildings on the left edge is emphasized while the importance of the buildings in the center is suppressed. The red area of the building in the right bottom corner is adjusted to an area with a similar shape of the building as marked by the left red box. The visualization in Fig. 12 indicates that the distance

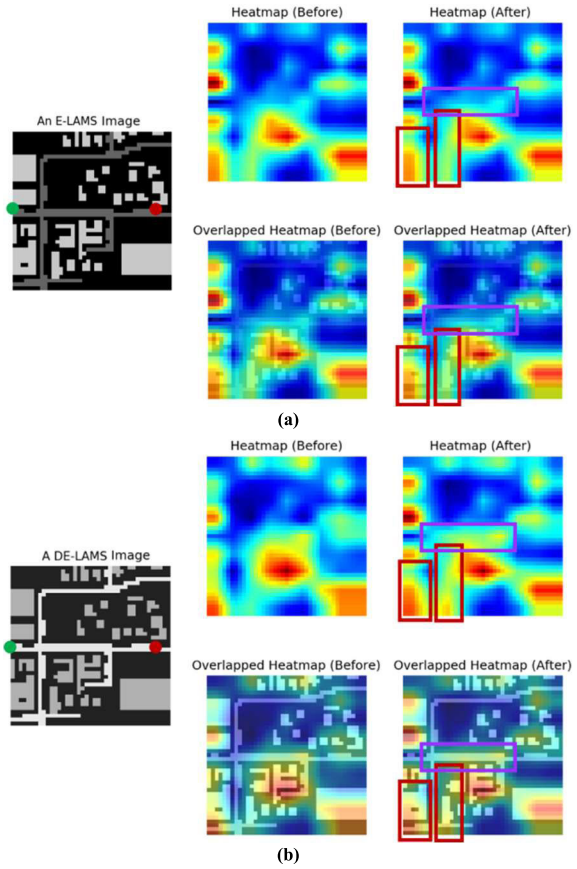


FIGURE 11. (a) An E-LAMS image and its overlapped heatmaps before and after using attention rendered by GC blocks. (b) A DE-LAMS image and its overlapped heatmaps before and after using attention. There are three columns with a total of five images in each figure. The image in the left column is the input image that is an E-LAMS image or a DE-LAMS image. The green and red points at the input image indicate the locations of Tx and Rx. The images in the center column are visualizations before using GC blocks, including a heatmap image, and the heatmap image overlapped on the input image. The images in the right column are visualizations after using GC blocks. The areas covered by red boxes and purple boxes contain emphasized buildings and streets by the attention, respectively.

information embedded in DE-LAMS images can help the AE-CNN model efficiently improve learning the latent features with dependency on distance. Besides, the visualization of heatmaps indicates that the global information on propagation environments is vital for mmWave path loss modeling using deep learning.

2) VISUALIZATION OF PREDICTED PATH LOSS

The AE-CNN model trained with DE-LAMS images achieved the best performance on each test scenario as presented in Table 4; its predicted path loss and the measured path loss in each test scenario are visualized in Figs. 12–17.

The performances of the proposed AE-CNN path loss model in S7 and S13 are shown in Figs. 12 and 13, respectively. The trend of predicted and measured path loss values can indicate the continuous variations of path loss with respect to the measurement index or the TR-separation. It can

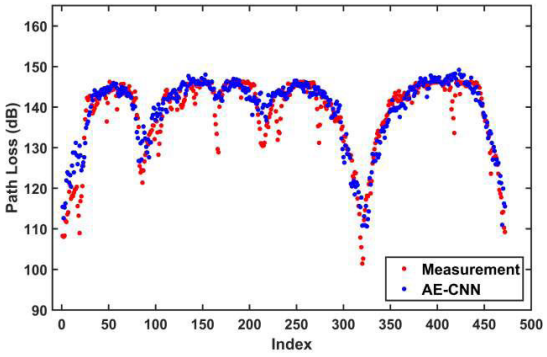


FIGURE 12. Measurement data and the predicted values of the AE-CNN model versus index in S7. The predicted path loss of the AE-CNN model and the measured path loss are denoted as "AE-CNN" and "Measurement", respectively.

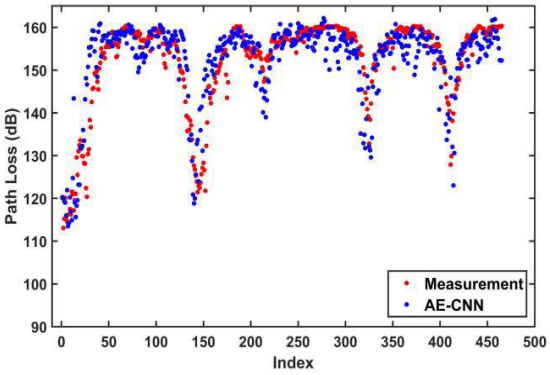


FIGURE 13. Measurement data and the predicted values of the AE-CNN model versus index in S13.

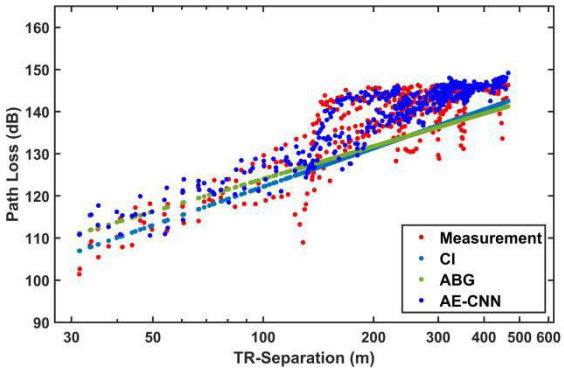


FIGURE 14. Measurement data and the predicted values of the CI and ABG model, and the predicted values of the AE-CNN model versus TR-separation in S7. The predicted path loss of the CI and ABG model are denoted as "CI" and "ABG", respectively.

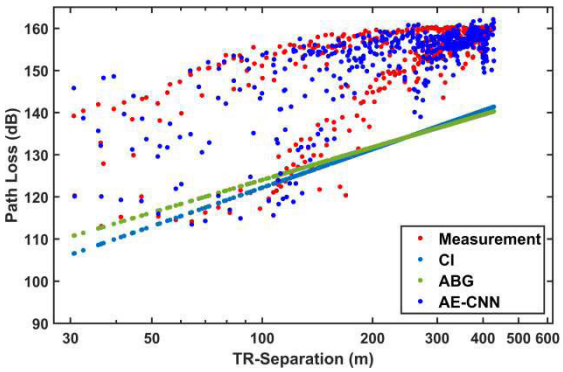


FIGURE 15. Measurement data and the predicted values of the CI and ABG model, and the predicted values of the AE-CNN model versus TR-separation in S13.

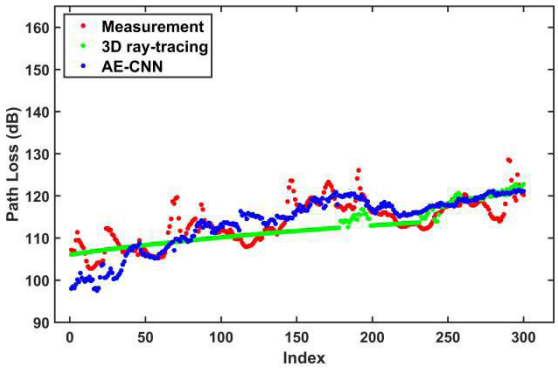


FIGURE 16. Measurement data and the simulated values of the 3D ray-tracing method, and the predicted values of the AE-CNN model versus index in S17. The simulated path loss of the 3D ray-tracing method is represented by "3D ray-tracing".

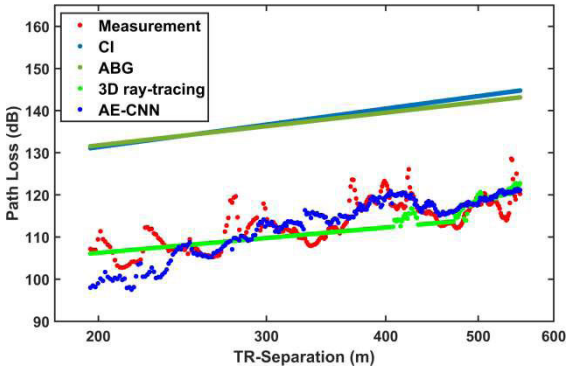


FIGURE 17. Measurement data and the predicted values of the CI and ABG model, simulated values of the 3D ray-tracing method, and the predicted values of the AE-CNN model versus TR-separation in S17.

be observed that the trend of the predicted path loss can match that of the measured path loss in S7 and S13. This indicates that the AE-CNN model can predict path loss values according to the given scenario and the specific Rx positions.

Figs. 14 and 15 show the performances of the proposed AE-CNN model with respect to the log-scale TR-separation in S7 and S13, respectively. As one can see, the predicted values of the AE-CNN model is close to the measurement data in both S7 and S13. The empirical models, including the CI and ABG model, show good performance when the TR-separation is smaller than 150 meters in S7, as shown in Fig. 14, but they underestimate the overall path loss when the TR-separation is higher than 150 meters. Fig. 15 shows the predicted path loss of the empirical methods is underestimated in S13. The empirical methods could not make

accurate predictions when the range of path loss values is changed because the transmission power of Tx varies in distinct scenarios. The proposed AE-CNN model can be immune from such changes.

Figs. 16 and 17 show the performance comparisons over the predicted values of the AE-CNN model and conventional methods. Compared with the simulated values of the 3D ray-tracing method, the predicted path loss of the proposed AE-CNN model achieves better performance with a real trend matching the measured path loss. Whereas the simulated path loss of the 3D ray-tracing method is almost distributed in a straight line related to log-scale TR-separation values, as shown in Fig. 16. This is caused by way of the deterministic methods to calculate path loss values. Since S17 is a scenario where all Rx points are LoS sites and its measurement data is recorded when a car with a rooftop Rx moved away from Tx, the influence of free space path loss is mainly considered for the 3D ray-tracing method. This phenomenon also exists in Fig. 17. The predicted values of empirical methods in Fig. 17 show an overall overestimation of the measured path loss. This indicates that the performance of empirical methods could be unsatisfactory when the applied scenarios are distinctly different from the scenarios used to calculate their parameters.

3) VISUALIZATION OF PREDICTED ERRORS

Fig. 18 shows the absolute predicted error of the CNN model and AE-CNN model with their polynomial fitting curves. The AE-CNN model shows the smaller absolute error in a wide range of the TR-separation than CNN model. The fitting curve of the AE-CNN model is flatter than that of the CNN model. This indicates that the proposed AE-CNN path loss model can achieve reliable performance in a large area. Further, the difference of absolute predicted error between the AE-CNN model and the CNN model increases with the decline of TR-separation. Besides, Fig. 3 (c) shows that the proposed AE-CNN model achieves better approximation with respect to the TR-separation in S7 by closer inspection. These observations also show the effectiveness of DE-LAMS images for the proposed model.

This section studied the influences of several approaches in the proposed AE-CNN path loss model on path loss prediction performance for 28 GHz mmWaves in suburban environments through extensive experiments. First, DE-LAMS images with the distance information improve learning the latent features with dependency on distance. Second, dilated convolution alleviates the locality of extracted features and captures the global information on input images. Third, the attention rendered by GC blocks utilizes global information from inputs to extract essential features having significant influences on path loss and suppress unnecessary features of propagation environments. With the captured global information via dilated convolution and attention from DE-LAMS images, the AE-CNN path loss model can extract features having significant influences on path loss and suppress unnecessary features.

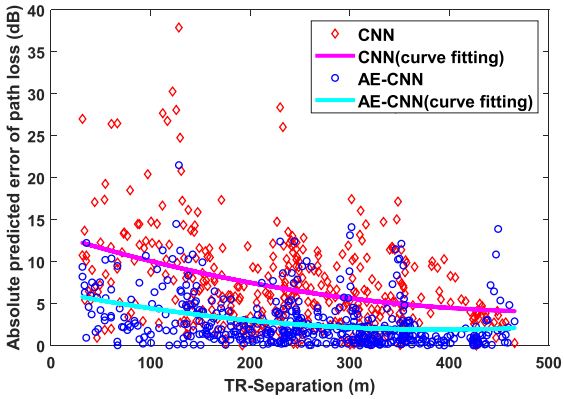


FIGURE 18. Predicted error of the CNN model and AE-CNN model versus TR-separation.

The comprehensive visualization with detailed analysis demonstrates the promising and satisfactory performance of the AE-CNN path loss model. Furthermore, the proposed AE-CNN path loss model does not require detailed information on the propagation environment, and it only needs to have access to 2D maps that are easy to obtain. The proposed model's input preprocessing procedures and generation of input images are pixel-level calculation and copying. They do not require a lot of computational resources and time. These advantages make the AE-CNN path loss model is more applicable in large-scale development.

V. CONCLUSION

This paper proposed a novel AE-CNN path loss model using deep learning with dilated convolution and attention. The use of the AE-CNN path loss model represents an efficient and straightforward approach to predict the path loss of 28 GHz mmWaves for 5G communications in suburban environments. The experimental results indicate the proposed AE-CNN path loss model can yield superior performance compared with popular conventional deterministic and empirical methods. The preprocessing of the 2D maps and the generation of input images is concise and efficient. The proposed AE-CNN path loss model can easily be applied to diverse scenarios due to its favorable advantages in practice.

Buildings and streets in suburban environments that significantly influence path loss are the main factors considered in this study for mmWaves. Other factors such as the influence of moving vehicles and the Tx and Rx height should be considered for further studies. We focus on the applying of deep learning technologies into path loss modeling in this study. However, the fine-tuning of the network structures and selecting more appropriate hyperparameters can be done with extensive experiments to improve the performance of the proposed AE-CNN model. Other state-of-the-art deep learning technologies should be considered on path loss modeling for 5G communications in suburban environments and urban environments.

ACKNOWLEDGMENT

The authors would like to thank Samsung Electronics for providing the measurement data and other assistance to the research in the work presented in this paper.

REFERENCES

[1] P. Cerwall, P. Jonsson, R. Möller, S. Bävertoft, S. Carson, and I. Godor, "Ericsson mobility report," On Pulse Netw. Soc. Hg. v. Ericsson, Stockholm, Sweden, Tech. Rep. EAB-20:009174, 2020. [Online]. Available: https://www.ericsson.com/4adc87/assets/local/mobility-report/documents/2020/ november-2020-ericsson-mobility-report.pdf

[2] F. Al-Ogaili and R. M. Shubair, "Millimeter-wave mobile communications for 5G: Challenges and opportunities," in *Proc. IEEE Int. Symp. Antennas Propag. (APSURSI)*, Jun. 2016, pp. 1003–1004.

[3] K. Haneda, J. Zhang, L. Tan, G. Liu, Y. Zheng, H. Asplund, J. Li, Y. Wang, D. Steer, and C. Li, "5G 3GPP-like channel models for outdoor urban microcellular and macrocellular environments," in *Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring)*, May 2016, pp. 1–7.

[4] Y. Zhang, S. Jyoti, C. R. Anderson, D. J. Love, N. Michelusi, A. Sprintson, and J. V. Krogmeier, "28-GHz channel measurements and modeling for suburban environments," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[5] J. Ko, Y.-J. Cho, S. Hur, T. Kim, J. Park, A. F. Molisch, K. Haneda, M. Peter, D.-J. Park, and D.-H. Cho, "Millimeter-wave channel measurements and analysis for statistical spatial channel model in in-building and urban environments at 28 GHz," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5853–5868, Sep. 2017.

[6] J. Du, D. Chizhik, R. Feick, G. Castro, M. Rodríguez, and R. A. Valenzuela, "Suburban residential building penetration loss at 28 GHz for fixed wireless access," *IEEE Wireless Commun. Lett.*, vol. 7, no. 6, pp. 890–893, Dec. 2018.

[7] Y. K. Yoon, K. W. Kim, and Y. J. Chong, "Site prediction model for the over rooftop path in a suburban environment at millimeter wave," *Int. J. Antennas Propag.*, vol. 2019, pp. 1–12, Apr. 2019.

[8] D. He, B. Ai, K. Guan, L. Wang, Z. Zhong, and T. Kurner, "The design and applications of high-performance ray-tracing simulation platform for 5G and beyond wireless communications: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 10–27, 1st Quart., 2019.

[9] K.-G. Lee, S.-J. Oh, J.-S. Woo, and K.-T. Lee, "Propagation characteristics of suburban environments using hybrid ray-tracing simulation," in *Proc. IEEE 84th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2016, pp. 1–5.

[10] I. Popescu, I. Nafornita, and P. Constantinou, "Comparison of neural network models for path loss prediction," in *Proc. IEEE Int. Conf. Wireless Mobile Comput., Netw. Commun.*, Aug. 2005, pp. 44–49.

[11] R. D. Timoteo, D. C. Cunha, and G. D. Cavalcanti, "A proposal for path loss prediction in urban environments using support vector regression," in *Proc. Adv. Int. Conf. Telecommun.*, 2014, pp. 1–5.

[12] B. J. Cavalcanti, G. A. Cavalcante, L. M. D. Mendonça, G. M. Cantanhede, M. M. M. D. Oliveira, and A. G. D'Assunção, "A hybrid path loss prediction model based on artificial neural networks using empirical models for LTE and LTE—A at 800 MHz and 2600 MHz," *J. Microw., Optoelectron. Electromagn. Appl.*, vol. 16, no. 3, pp. 708–722, Sep. 2017.

[13] X. Zhao, C. Hou, and Q. Wang, "A new SVM-based modeling method of cabin path loss prediction," *Int. J. Antennas Propag.*, vol. 2013, pp. 1–7, May 2013.

[14] S. Yang and H. Lee, "Feature extraction for neural network wave propagation loss models from field measurements and digital elevation map," *IEICE Trans. Electron.*, vol. 82, no. 7, pp. 1260–1266, 1999.

[15] S. I. Popoola, E. Adetiba, A. A. Atayero, N. Faruk, and C. T. Calafate, "Optimal model for path loss predictions using feed-forward neural networks," *Cogent Eng.*, vol. 5, no. 1, Jan. 2018, Art. no. 1444345.

[16] E. Ostlin, H. Zepernick, and H. Suzuki, "Macrocell path-loss prediction using artificial neural networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2735–2747, Jul. 2010.

[17] Y. Zhang, J. Wen, G. Yang, Z. He, and X. Luo, "Air-to-air path loss prediction based on machine learning methods in urban environments," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–9, Jun. 2018.

[18] Y. Zhang, J. Wen, G. Yang, Z. He, and J. Wang, "Path loss prediction based on machine learning: Principle, method, and data expansion," *Appl. Sci.*, vol. 9, no. 9, p. 1908, May 2019.

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

[20] S. Hur, S. Baek, B. Kim, Y. Chang, A. F. Molisch, T. S. Rappaport, K. Haneda, and J. Park, "Proposal on millimeter-wave channel modeling for 5G cellular system," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 454–469, Apr. 2016.

[21] S. P. Sotiroudis, S. K. Goudos, and K. Siakavara, "Deep learning for radio propagation: Using image-driven regression to estimate path loss in urban areas," *ICT Exp.*, vol. 6, no. 3, pp. 160–165, Sep. 2020.

[22] S. Moon, H. Kim, and I. Hwang, "Deep learning-based channel estimation and tracking for millimeter-wave vehicular communications," *J. Commun. Netw.*, vol. 22, no. 3, pp. 177–184, Jun. 2020.

[23] H. Cheng, S. Ma, and H. Lee, "CNN-based mmWave path loss modeling for fixed wireless access in suburban scenarios," *IEEE Antennas Wireless Propag. Lett.*, vol. 19, no. 10, pp. 1694–1698, Oct. 2020.

[24] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.

[25] A. Galassi, M. Lippi, P. Torroni, and L. Systems, "Attention in natural language processing," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Sep. 10, 2020, doi: 10.1109/TNNLS.2020.3019893.

[26] H. Fukui, T. Hirakawa, T. Yamashita, and H. Fujiyoshi, "Attention branch network: Learning of attention mechanism for visual explanation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 10705–10714.

[27] W. Wang and J. Shen, "Deep visual attention prediction," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2368–2378, May 2018.

[28] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5998–6008.

[29] F. Wang and D. M. J. Tax, "Survey on the attention based RNN model and its applications in computer vision," 2016, *arXiv:1601.06823*. [Online]. Available: http://arxiv.org/abs/1601.06823

[30] K. Cho, A. Courville, and Y. Bengio, "Describing multimedia content using attention-based encoder-decoder networks," *IEEE Trans. Multimedia*, vol. 17, no. 11, pp. 1875–1886, Nov. 2015.

[31] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," 2015, *arXiv:1511.07122*. [Online]. Available: http://arxiv.org/abs/1511.07122

[32] Y. Cao, J. Xu, S. Lin, F. Wei, and H. Hu, "GCNet: Non-local networks meet squeeze-excitation networks and beyond," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 1971–1980.

[33] N. Parmar, A. Vaswani, J. Uszkoreit, L. Kaiser, N. Shazeer, A. Ku, and D. Tran, "Image transformer," 2018, *arXiv:1802.05751*. [Online]. Available: http://arxiv.org/abs/1802.05751

[34] J. Hu, L. Shen, and G. Sun, "Squeeze- and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.

[35] L. Chen, H. Zhang, J. Xiao, L. Nie, J. Shao, W. Liu, and T.-S. Chua, "SCA-CNN: Spatial and channel-wise attention in convolutional networks for image captioning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 5659–5667.

[36] S. Woo, J. Park, J.-Y. Lee, and I. So Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 3–19.

[37] R. W. Jones, "Spectrum monitoring handbook," in *Proc. Int. Telecommun. Union*, 2002, pp. 262–362.

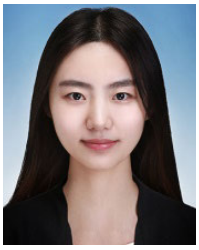
[38] P. Newson and J. Krumm, "Hidden Markov map matching through noise and sparseness," in *Proc. 17th ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, 2009, pp. 336–343.

[39] E. Korsberg and E. Nordén, "Lane-level map matching using hidden Markov models," M.S. thesis, Dept. Mech. Maritime Sci., Chalmers Univ. Technol., Göthenburg, Sweden, 2019. [Online]. Available: https://publications.lib.chalmers.se/records/fulltext/256993/256993.pdf

[40] J. Yao, A. Tabatabaei Balaei, M. Hassan, N. Alam, and A. G. Dempster, "Improving cooperative positioning for vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2810–2823, Jul. 2011.

[41] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, and M. Isard, "TensorFlow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Operating Syst. Design Implement.*, 2016, pp. 265–283.

HONG CHENG was born in Qingdao, Shandong, China, in 1993. He received the B.S. degree from the Department of Internet of Things Engineering, Qingdao University of Science and Technology, in 2016. He is currently pursuing the Ph.D. degree in computer engineering with Kwangwoon University, Seoul, South Korea. His research interests include deep learning and path loss modeling.



SHENGJIE MA was born in Linyi, Shandong, China, in 1994. She received the B.S. degree from the Department of Software Engineering, Qingdao University of Science and Technology, in 2016. She is currently pursuing the Ph.D. degree in computer engineering with Kwangwoon University, Seoul, South Korea. Her research interests include deep learning and path loss modeling.

HYUKJOON LEE was born in Seoul, South Korea, in 1963. He received the B.S. degree in computer science from the University of Michigan, Ann Arbor, USA, in 1987, and the M.S. and Ph.D. degrees in computer and information science from Syracuse University, in 1989 and 1993, respectively.

Since 1994, he has been working as a Senior Researcher with Samsung Electronics. In 1996, he joined the Faculty of the Department of Computer Engineering, Kwangwoon University. His research interests include machine learning, mobile and wireless networks, ITS, and smart healthcare systems.

MINSUNG CHO received the M.S. degree in electrical and telecommunication engineering from Sungkyunkwan University, South Korea, in 2000.

He currently works as a Principal Research Engineer with the Network Division, Samsung Electronics Company Ltd. His research interests include RAN architecture, RAN network design and optimization, radio planning tool design, and air interface analysis.