

AUTHOR QUERIES

PLEASE NOTE: We cannot accept new source files as corrections for your article. If possible, please annotate the PDF proof we have sent you with your corrections and upload it via the Author Gateway. Alternatively, you may send us your corrections in list format. You may also upload revised graphics via the Author Gateway.

Carefully check the page proofs (and coordinate with all authors); additional changes or updates **WILL NOT** be accepted after the article is published online/print in its final form. Please check author names and affiliations, funding, as well as the overall article for any errors prior to sending in your author proof corrections.

AQ:1 = The abbreviation FAM has been used for the terms “feature activation map” and “feature activation mapping” Please confirm which has to be followed.

AQ:2 = Please confirm or add details for any funding or financial support for the research of this article.

AQ:3 = Please provide the full current affiliation details (department name, name of university/institution, city, country, and postal code) for all the authors.

AQ:4 = The in-text citations of “Fig. 2” is out of order. Fig. 1 should be cited before Fig. 2. Please update the in-text citations.

AQ:5 = Please confirm the retention of the sentence in the acknowledgment section is correct.

AQ:6 = Please provide the accessed date for Refs. [11], [27], [28], [29], [30], and [31].

Deep Learning-Based THz Wireless Channel Property Prediction in Motherboard Desktop Environment

Jinbang Fu^{ID}, Erik J. Jorgensen^{ID}, *Graduate Student Member, IEEE*, Prateek Juyal^{ID},
and Alenka Zajić^{ID}, *Senior Member, IEEE*

Abstract—This article proposes a residual network (ResNet)-based feature concatenated neural network model to predict the type of scenario the channel is under and the attribute of the predicted scenario with power delay profile (PDP) as the inputs. The generalized model structure consists of three blocks for feature extraction, scenario prediction, and attribute prediction, respectively. The PDP data is collected from a motherboard desktop environment under five different physical arrangement scenarios. Within each scenario, data is collected several times while varying a different physical attributes for each scenario. Two steps of data augmentation are applied to expand the size and to improve the resolution (difference between the neighboring attributes) of the measured dataset for the robust training and thorough evaluation of the proposed model. The proposed model is evaluated and compared with a multilayer perceptron (MLP)-based model on an expanded measured and averaged interpolated dataset. It is shown that both models perform very well on the expanded measured dataset with nearly 100% prediction accuracy on both scenarios and attributes. The MLP-based model suffers performance degradation on the averaged interpolated dataset with up to a 9% drop of classification accuracy on attribute prediction tasks, while our ResNet-based feature concatenated model performs equally in both scenarios. **Feature activation map (FAM)** and grad-class activation mapping (Grad-CAM) approaches are applied to provide visual explanations highlighting characteristics of the input PDP used for model decisions. FAM shows that the MLP-based model focuses on the multipath generated peaks of the PDP where some interpolated neighboring data points cannot be distinguished. The Grad-CAM shows that the proposed ResNet-based feature concatenated model performs better because it has strong attention not only on the multipath peaks but also on the valleys between those peaks which hold distinguishing information.

Index Terms—Channel characterization, channel prediction, channel sounding, chip-to-chip wireless channels, THz communications.

I. INTRODUCTION

THE THz wireless link is considered as a replacement for wired connections in inter-chip communication for future computing systems [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]. A wireless link system alleviates the needs for

cable management and reduces the complexity of the computing system design [1], [2]. Operating at THz frequencies with its larger available bandwidth allows higher speed and lower latency communications. With smaller antenna sizes and reduced spacing at THz frequencies, spatial multiplexing can be applied to improve data rates to Terabit-per-second which is close to the data rate of current wired systems [2], [11]. To realize THz wireless inter-chip communication, comprehensive understanding, and identification of the corresponding wireless channels is required. Wireless channels are usually determined by the signal propagation environment and interactions with the objects in the environment, also known as the channel properties [12]. These channel properties define the characteristics of the wireless channel and affect the quality of possible wireless links. The prediction of channel properties is important for establishing inter-chip communication and allows for further adjustments to improve the communication link. With a precise prediction of channel properties, the appropriate transmission mode can be selected to adapt to the environment, which can help to provide more reliable communication with better performance [12].

To realize THz wireless inter-chip communication, channel modeling tools are needed and several THz wireless channel models have been proposed [8], [9], [13], [14], [15]. Based on the geometrical distribution of the environment, statistical channel models have been proposed for short-range device-to-device channels and THz wireless inter-chip communication channels [3], [4]. Constant rate models have been applied on THz wireless channels in a datacenter and in the motherboard desktop environment by modifying the classic Saleh-Valenzuela (S-V) model [6], [16]. These classic analytical techniques can provide good estimations of the wireless channel with simple math equations. However, these techniques are generally not suitable for the prediction and identification of channel properties because the simplifications introduced in the models may not capture more complex multipath environment properties [17]. Moreover, the large number of potential links, obstructions, scattering, and so on for the inter-chip communication in motherboard desktop environment requires a large amount of data for a thorough channel characterization. These classical techniques are less suitable for the analysis of large datasets as they might miss important relations within the data [18].

Manuscript received 16 June 2022; revised 20 March 2023; accepted 14 May 2023. This work was supported in part by the NSF under Grant 1651273. (Corresponding author: Jinbang Fu.)

The authors are with ??? (e-mail: jfu72@gatech.edu).
Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TAP.2023.3278831>.

Digital Object Identifier 10.1109/TAP.2023.3278831

In contrast, machine learning (ML) is a branch of artificial intelligence that learns a model from data to make decisions and has achieved breakthroughs in application areas such as image processing, natural language processing, and data mining [18]. Due to their rise in popularity, ML techniques have been gradually applied to study the wireless propagation channel, such as channel characterization and channel modeling [10], [12]. Researchers have applied ML techniques on the estimation of propagation loss [17], [19], tracking multipath components [20], and derivation of channel statistical parameters [21]. Also, ML techniques are applied on channel property prediction as they have been shown to provide good performance in identifying small differences between measurements to characterize different properties. Several deep learning-based methods have been proposed for a nonline of sight (NLoS) recognition [22], indoor environments classification [23], and vehicle communication scenarios identification [24]. Considering the complicated propagation channel on the densely populated motherboard in a metal casing, where some hidden patterns due to scattering, reflections, and resonant cavity effects are difficult to identify by traditional techniques, ML methods are more likely to succeed for channel prediction in such an environment.

To establish a stable wireless link for THz wireless inter-chip communication in computing systems, it is necessary to understand the channel properties in advance. These properties include potential physical scenarios of the channel such as the link between CPU and PCI, the link between CPU and dual in-line memory module (DIMM), and so on, and certain attributes under those scenarios such as the height differences between the chips or the distances between the CPU and the PCI components. These channel properties could be predetermined by an ML model with the information given by the pilot signal.

The main contributions of this article are summarized as follows.

- 1) We apply deep learning methods to predict channel properties for THz chip-to-chip wireless communication in motherboard desktop environment. With PDP as the input, we propose a residual network (ResNet)-based feature concatenated model to predict the channel's physical scenario and an additional attribute of that predicted scenario.
- 2) The proposed model is evaluated and compared with a multilayer perceptron (MLP)-based model on the expanded measured dataset and the averaged interpolated dataset. Results show that our model performs very well on the expanded measured dataset with nearly perfect prediction of both scenarios and attributes. These results exemplify the validity of our proposed model and the feasibility of applying deep learning model for THz channel prediction on a densely-populated motherboard in a practical computing system. We show that our ResNet-based feature concatenated model performs more robustly than the MLP model as it has higher classification accuracy when predicting a specific scenario's attribute. Our ResNet-based feature concatenated model also has constant performance on both measured

and augmented dataset while the MLP model shows performance degradation on the averaged interpolated dataset.

- 3) feature activation mapping (FAM) and Gradient-weighted class activation mapping (Grad-CAM) are applied on the MLP and our ResNet-based feature concatenated models, respectively, to explain their performance differences by determining the contributing regions of the raw data that the models use to make their predictions [25], [26]. The results provide us a deeper understanding of the THz wireless channel in the motherboard desktop environment. We find that the MLP model only focuses on the peaks of the input PDP with the delay of 0, 1.835, 3.837 ns corresponding to multipath propagations, while the ResNet-based feature concatenated model has strong attention not only to those peaks but also on the valleys between those peaks. This implies that the valleys between the peaks generated by the multipath signals also contain important information related to the wireless channel that help distinguish different channel properties.

The rest of the article is organized as follows. Section II presents the channel measurements and the augmentation of the collected data. Section III illustrates the design of the ResNet-based feature concatenated neural network model. Section IV presents the analysis and model verification. Finally, Section V makes concluding remarks.

II. DATA MEASUREMENTS AND AUGMENTATIONS

In this section, we provide a brief discussion about THz wireless channel measurements in the motherboard desktop environment under five different potential scenarios: LoS, reflected non-LoS (RNLoS), obstructed-LoS (OLoS), non-LoS (NLoS), and a practical CPU-PCI link. Under each of these scenarios, measurements are performed several times while varying a different attribute for each scenario. Due to time and physical setup limitations, we also apply multiple data augmentation techniques on the measured data for training, validation, and analysis of our proposed models.

A. THz Channel Measurements

The measurement setup includes a Keysight N5224A PNA vector network analyzer (VNA), a Virginia diodes, Inc. (VDI) transmitter (Tx210), and a VDI receiver (Rx148) [27]. The VNA generates the input signal in the range of 10 MHz–12 GHz with the power level (P_{in}) of 0 dBm and feeds it into the VDI Tx210. In the VDI Tx210 transmitter, a 25 GHz signal is generated by a Herley-CTI phase-locked dielectric resonator oscillator (DPRO with 100 MHz reference crystal oscillator) [28]. It is amplified and frequency doubled by Norden N08-1975 [29] and then tripled by VDI WR6.5X3 [30]. This signal is then fed into a sub-harmonic mixer (WR2.8SHM) [31] which doubles the carrier frequency and mixes it with the VNA-generated baseband signal. This THz-range signal is then transmitted by a horn antenna in the range of 288–312 GHz. The same components are used to down-convert the signal, the difference is that the DPRO is

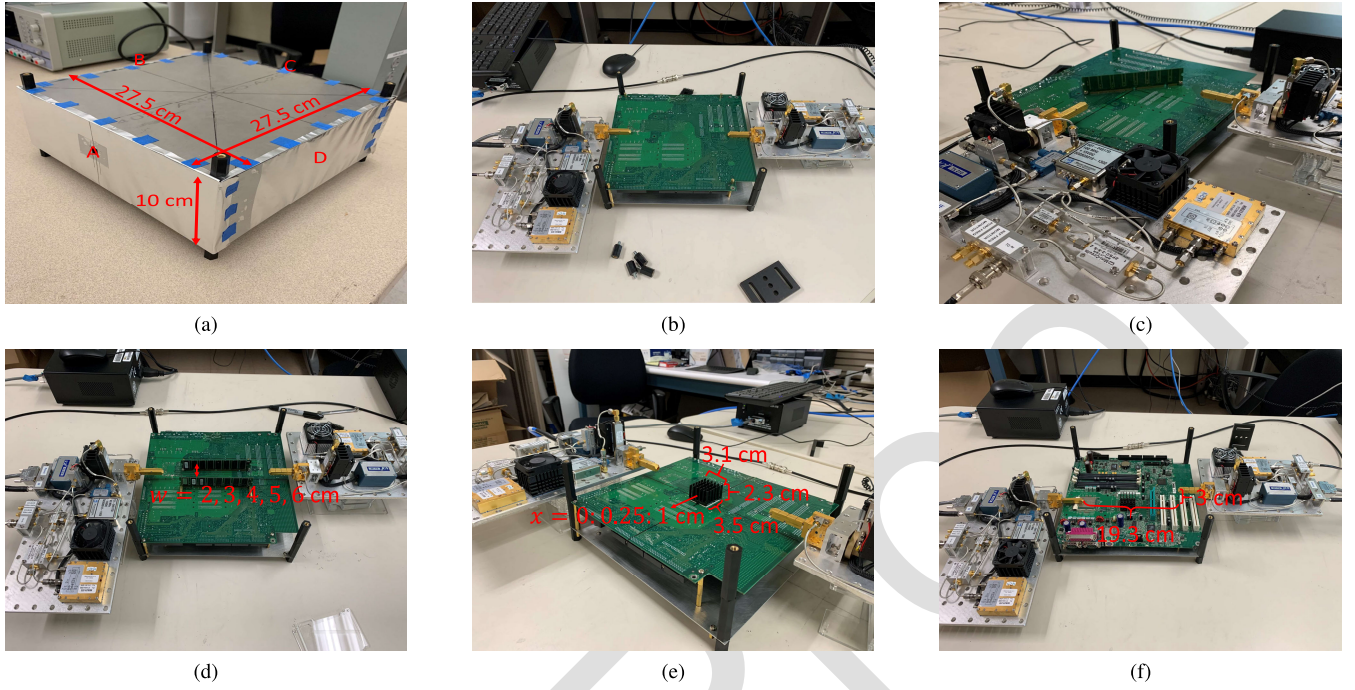


Fig. 1. Setups for the measurements of THz wireless channel in motherboard desktop environment. (a) Fabricated metal casing that houses each of (b) LoS, (c) RNLoS, (d) OLoS, (e) NLoS, and (f) CPU-PCI link measurement scenarios.

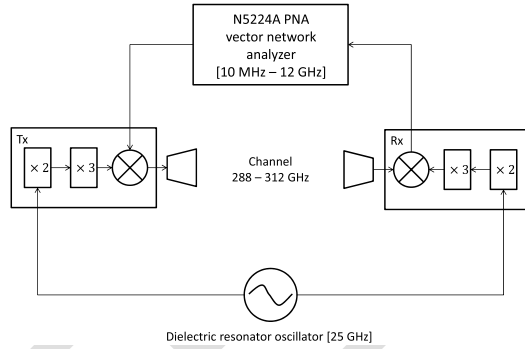


Fig. 2. Block diagram of the measurement system.

TABLE I
MEASUREMENT PARAMETERS

Parameter	Symbol	Value
Measurement points	N	801
Intermediate frequency bandwidth	Δf_{IF}	20 kHz
Average noise floor	P_N	-90 dBm
Input signal power	P_{in}	0 dBm
Start frequency	f_{start}	10 MHz
Stop frequency	f_{stop}	12 GHz
Bandwidth	B	11.99 GHz
Time domain resolution	Δt	0.083 ns
Maximum excess delay	τ_m	33.4 ns

tuned to 24.2 GHz. The received signal is then down-converted to an intermediate frequency (IF) of 9.6 GHz. The VNA then samples the upper sideband of the down-converted signal in the range of 9.61-21.6 GHz with 801 points. The block diagram of the measurement system is shown in Fig. 2 [1], [32] and all measurement parameters are summarized in Table I. Pyramid horn antennas with a gain that varies in the range of 22–23 dBi over the observed frequency band from 300 to 312 GHz were used in the measurements. The theoretical half-power beamwidth (HPBW) of the horn is about 12° in azimuth and elevation. The physical dimension of the horn aperture is 8.91 mm, which limits the far-field boundary to 15.88 cm at 300 GHz according to the Fraunhofer distance.

To measure the THz wireless channel in a motherboard desktop environment, we fabricated an aluminum metal cavity with the size of $27.5 \times 27.5 \times 10$ cm, which approximates the size of a desktop, with two square aluminum plates as the top and bottom walls [5]. The other four sides are

wrapped by aluminum foil and labeled as A–D, as shown in Fig. 1(a). As shown in Fig. 1(b), a motherboard was supported 4.2 cm over the bottom wall by brass hex standoffs inside the fabricated cases. For all channel environment scenarios, measurements are performed on the motherboard inside the metal cavity.

For the measurements of LoS links, as shown in Fig. 1(b), antennas of the transmitter (Tx) and receiver (Rx) were aligned with each other and inserted into the cavity through small holes drilled on the transceiver sidewalls (A and C) of the cavity. Measurements were performed on the backside of the motherboard to separate the influence of the other components. As the heights of components on the motherboard may differ and to investigate the ground effect of the motherboard, both

transceivers were elevated simultaneously in height during the measurements. To do this, the height of transceivers h was varied from 0 to 4.5 cm above the motherboard in increments of 0.3 cm while the distance between the transceivers was fixed at 19.3 cm.

As shown in Fig. 1(c), a DIMM was placed standing up on the center of the motherboard and diagonally aligned with transceivers which were positioned orthogonal to each other for the RNLoS measurements. The distance from the DIMM to each of the transceiver is 9.55 cm and the heights for both transceivers were fixed at 1.8 cm above the motherboard. Both flat and component sides of the DIMM were investigated during the measurements as both sides could possibly be used to establish the wireless link on the motherboard. For the OLoS link, measurements were performed to investigate the effects of parallel-plate structures on the motherboard. As shown in Fig. 1(d), two DIMMs were put in parallel on the center of the motherboard. when the distance between parallel plates, ω , is small, e.g., $\omega = 2$ or 3 cm, some parts of the parallel plates may be in the second or third region of the Fresnel ellipsoid space whose radius is 1.17 or 1.44 cm. The interceptions may lead to destructive or constructive effects due to the phase shift of the signal reflecting on the parallel plates [2]. Measurements were performed by varying the distances between the DIMMs from 2 to 6 cm with the step increment of 1 cm to simulate the variable distances between parallel structures on the motherboard. The heights of both transceivers were kept at 1.8 cm during the measurements.

The NLoS link was measured with a heatsink of size $3.5 \times 3.1 \times 1.3$ cm as an obstacle, as shown in Fig. 1(e). Between measurements, the heatsink was horizontally shifted away from the transceiver toward side D of the cavity, with a step increment of 0.25 cm to understand how the wireless link is affected by the gaps between the fins of the heatsink. The distance from the heatsink to each of the transceiver is 8.1 cm and the heights of both transceivers were kept at 1.8 cm.

Besides the investigations of these primary controlled scenarios, a practical CPU-PCI link was also measured on the densely populated motherboard. As shown in Fig. 1(f), Tx was mounted on top of the CPU and Rx was hung over the PCI slot with a 3 cm T-R vertical misalignment, which approximates the height difference between the CPU and the chip on the peripheral components. During the measurements, the horizontal T-R separations were gradually increased from 19.3 to 23.3 cm with the step increment of 1 cm as there are several PCI slots on the motherboard with different distances to the CPU. All the scenarios and corresponding variable attributes are summarized in Table II.

B. Data Augmentations

To robustly train our deep neural network model to precisely predict channel properties, we require a large dataset to feed into the model. However, measurements inside the metal cavity are very time-consuming as we need to rebuild the cavity every time we shift the transceivers or components for the measured attribute of each scenario. Additionally, due to physical constraints, we choose a relatively large step increment (low resolution) for the measurements under each

TABLE II
MEASURED POTENTIAL SCENARIOS AND CORRESPONDING ATTRIBUTES

LoS	The heights of transceivers
RNLoS	Flat/component side of the DIMM
OLoS	The distance between parallel DIMMs
NLoS	The movement of the heatsink
CPU-PCI link	The distance between the transceivers

scenario to ensure we can physically capture the data precisely. For instance, the heights of transceivers were varied during the LoS measurements by stacking plastic plates with a thickness of 0.3 cm under the transceivers. This means that information between the measured points is unknown to us. To validate the generalization ability of our proposed model, we also test them on a dataset with higher resolution. To that end, we apply data augmentation to expand the size of our dataset and increase its resolution by creating synthetic interpolated data from measured results. The tweaks of training examples in data augmentation make the model less likely to overfit on certain patterns, which improves the generalization ability of the model.

The power delay profile (PDP), which provides the intensity of a signal received through a multipath channel with respect to time delay, is fed into our deep neural network model for training and prediction. Measured PDP can be derived by squaring the inverse discrete Fourier transform (IDFT) of the measured frequency response, which itself is calculated by subtracting the S21 from the sum of antenna gains and the thru loss. Before any S21 measurements, the thru loss was first measured on the transceiver devices ten times. Then, as mentioned in II-A, we measured the S21 ten times at different attributes under each potential scenario. By combining all pairs of the ten thru loss and ten S21 measurements, we generate 100 samples for each attribute, as shown in Fig. 3(a). The expanded measured dataset was then separated into three groups by randomly separating 50%, 20%, and 30% of the data at each scenario and attribute to training, validation, and testing sets, respectively. Then the training and validation datasets were interpolated to a higher resolution for the scenarios which are measured in increments (all except the RNLoS scenario). For this augmentation, we applied linear, spline (spline of order 1), quadratic, cubic, PCHIP (piecewise cubic Hermite interpolating polynomial), and Akima interpolations to increase resolution down to steps of 0.1 cm (1λ) for LoS, OLoS, and CPU-PCI link measurements and to 0.05 cm (half λ) for the NLoS measurement. Interpolations were not applied on the data of RNLoS link because it is a binary dataset showing the difference between the flat or component side of the DIMM as the reflecting surface. Finally, we generated the average augmentation dataset by averaging the six interpolations for each attribute since we do not necessarily know which interpolation technique would be represent the unmeasured data. For example, data was measured at $h = 0, 0.3, 0.6, \dots, 4.5$ cm for the LoS propagation. Measurements were performed ten times at each height.

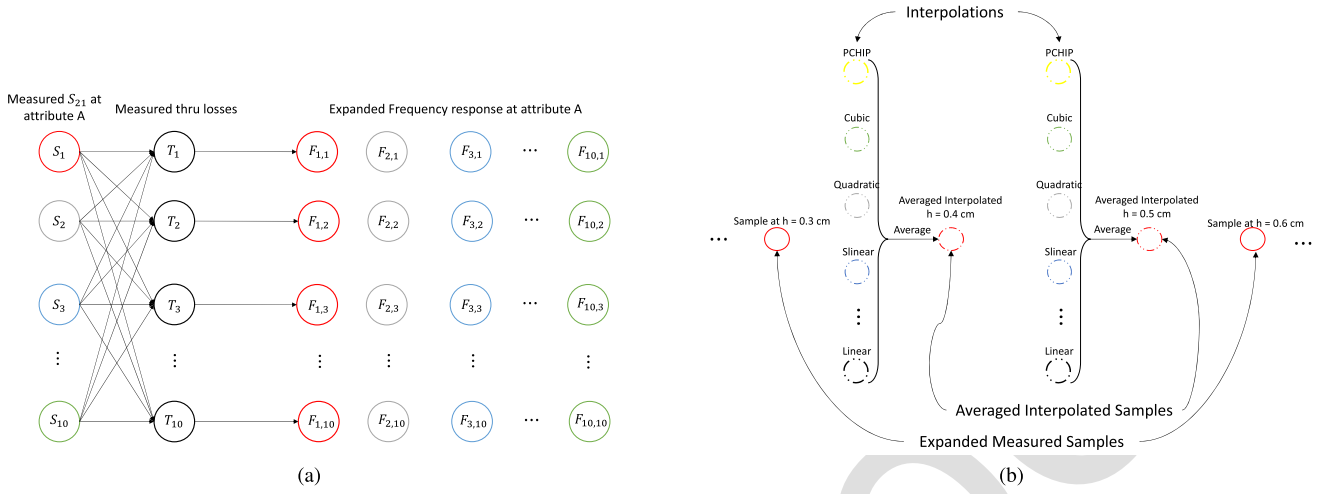


Fig. 3. Illustration of data augmentation: (a) data expansion and (b) averaged interpolation for data of LoS propagation.

With the data expansion, we generated 100 PDP samples at each height. As shown in Fig. 3(b), six different interpolation techniques were then applied to the training and validation datasets to improve the resolution from 0.3 to 0.1 cm. We then averaged those interpolations at each height. Doing so resulted in 50 and 30 PDP at each height $h = 0, 0.1, \dots, 4.5$ cm for the respective training and validation datasets, and 20 PDP at each of the same heights for the testing dataset. This averaged augmentation provides a single dataset (only one average PDP instead of six at each interpolated attribute) with higher resolution.

III. DEEP LEARNING MODEL

In this section, we introduce a ResNet-based feature concatenated neural network which is designed to predict properties of the THz wireless channel in the motherboard desktop environment. The general structure of the model is shown in Fig. 4. As shown in the figure, the proposed neural network model consists of three main blocks. The first is the “feature extraction” block which extracts the generalized features necessary to distinguish between scenarios and specific attributes of each scenario from the input PDP data. Those extracted features are then sent to the “scenarios prediction” block to predict the type of scenario the channel is under (LoS, RNLoS, OLoS, NLoS, and CPU-PCI link). The overall network design is called “feature concatenated” because the model’s scenario prediction output is then concatenated with the extracted features and fed to the “attributes prediction” block to estimate the attribute of that predicted scenario.

Our novel ResNet-based feature concatenated neural network model for the prediction of THz channel property is shown in Fig. 5. As shown in the figure, the feature extraction block of the proposed ResNet-based feature concatenated network model contains five sub-blocks, which are **b1**, **b2**, **b3**, **b4**, and a global average pooling (GAP) sub-block [33]. **b1** consists of a 64-channel 1×7 convolutional layer with the padding of 3 and a stride of 2, a batch normalization (BN) layer, rectified linear unit (ReLU) activation layer, and a 1×3 maximum pooling layer with a stride of 2. During

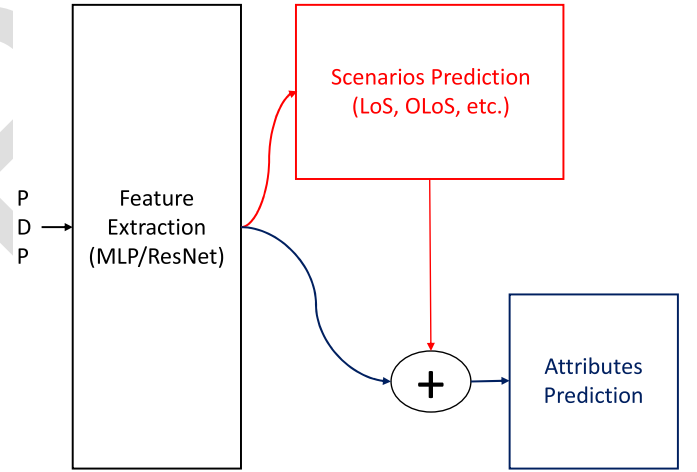


Fig. 4. Generalized model structure consisting of three main blocks.

training, the weights of intermediate convolution filters (e.g., the convolutional layers) may have strong variation between different layers, nodes in the same layer, and over time due to the updating model parameters. BN [34] is implemented to maintain stable optimization and reduce these strong variations that may hinder the convergence of the network. The ReLU [35] activation function is applied to increase the nonlinear expressiveness of the network, which simply outputs its input for positive values and 0 for negative values. A max-pooling layer is used to mitigate sensitivity of the network to noisy inputs and to spatially downsample representations [36]. **b1** increases the number of channels of the inputs from 1 to 64 while reducing the size of the data from 400 to 99. **b2–b4** are three ResNet building blocks shared with the same structures. Each of them consists of a residual block with an additional 1×1 convolutional layer to transform the inputs to the desired shape for the addition operation, and a residual block without the 1×1 convolutional layer. A residual block is the core element of a ResNet model, which guarantees the strictly increasing expressiveness of the network. The structure

multiple scenarios were nonzero. Similar to the structure of the scenario prediction block, the attribute prediction block also consists of three fully connected layers whose input sizes are 256, 128, and 32. In contrast with the scenario prediction block, the attribute of the scenario is predicted directly by the block.

Since the proposed network model predicts both the measurement scenario and an attribute of that scenario, training requires both a scenario classification loss and an attribute regression loss. Hence, the overall training loss function of the proposed model is derived as the sum of the cross-entropy loss for the scenarios classification and the mean squared error (MSE) loss for the attribute regression. The cross-entropy loss calculates the expected loss value for the predicted scenario class, while MSE quantifies the distance between the real and predicted value of the attribute. To train the network, we use the robust “Adam” variant of the stochastic gradient descent optimization algorithm which estimates the first- and second-order moments of the gradient via an exponential moving average to update network parameters [38].

IV. ANALYSIS AND MODEL VERIFICATION

In this section, we first trained the proposed ResNet-based feature concatenated network model described in Section III on the augmented training PDP. We then tested and compared the performance of the proposed model with an MLP-based network model on both the expanded measured dataset (without interpolations) and the averaged interpolated dataset (with interpolations) described in Section II-B. Finally, the ResNet- and MLP-based feature concatenated network models are analyzed with Grad-CAM and FAM, respectively, which provide the visual explanations for model decisions.

A. Model Training and Performance Evaluation

Before training, we first standardize the dataset by subtracting the mean and dividing by the standard deviation across the feature dimension. We also normalize the attributes for each scenario. These normalizations allow the network to learn weights on a similar scale for each scenario and attribute. The ResNet-based feature concatenated network model was trained for 1500 epochs. During each epoch, the network iterates over random batches of 256 samples and updates the variables in convolutional and linear layers with each batch. The loss of the model on training, validation, and testing dataset during the training process are compared in Fig. 7. It can be seen from the plot that the Loss of the model on the training dataset is always lower than the loss on the validation and testing dataset. The losses dropped quickly in the first 400 epochs of training as the model parameters are initialized randomly. The losses on the training, validation, and testing datasets converged to 5×10^{-5} , 6.5×10^{-4} , and 5.8×10^{-4} after 1200 epochs of training.

To evaluate the performance of the proposed ResNet-based feature concatenated network model, we first tested it on the expanded measured dataset. The results show that the model can perfectly classify the scenarios under which the input PDP was measured. As described in Section III, our proposed model

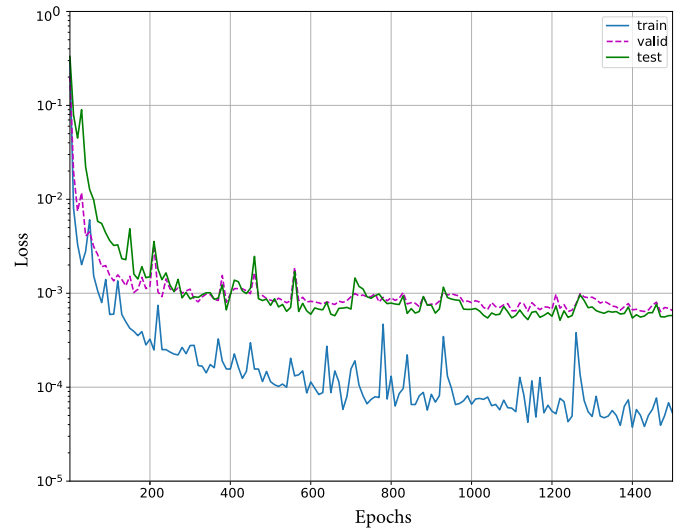


Fig. 7. Losses of ResNet-based feature concatenated network model on training, validation, and testing dataset over 1500 epochs of training.

predicts the attribute under the predicted scenario directly. As with any regression problem, there is always a deviation between the predicted and the correct value. For a robust model, that deviation is small so that the predicted value can be classified as the correct value by choosing an appropriate threshold. To evaluate the robustness of our proposed model, we calculated and compared the classification accuracy of the predicted attribute for each potential scenario with respect to a threshold in the range of zero to half of the step increment size set for the measurements. The maximum distance threshold is 0.15, 0.5, 0.5, 0.125, and 0.5 cm for LoS, RNLoS, OLoS, NLoS, and CPU-PCI link, respectively, as shown in Fig. 8(a). It can be seen from the plot that the ResNet-based feature concatenated model performs perfectly on the expanded measured dataset with achieved nearly 100% classification accuracy under all potential scenarios with the largest possible threshold. Also, it can be observed that the model differentiates between classes most easily in the RNLoS scenario as there are only two classes (either flat or component side of the DIMM as the reflection surface). Since there are only 5 different attributes for the NLoS and CPU-PCI link scenarios, the model differentiates between classes even with relatively small thresholds. Also, comparing these two scenarios, the model converges slower under the CPU-PCI link as the environment of this scenario is much more complicated than that of the NLoS. The small distances between the 16 different attributes for LoS result in slightly less robust performance across a range of thresholds. The model converges the slowest under the OLoS scenario as the measured PDP are very similar to each other with only small variations on the intensity of multipath signals.

The performance of our proposed ResNet-based feature concatenated network model has been compared with an MLP-based method, which is a widely used channel prediction scheme [12], [17], [21], [39], [40], [41], [42], [43], [44], [45], [46], [47]. The structure of the MLP-based network

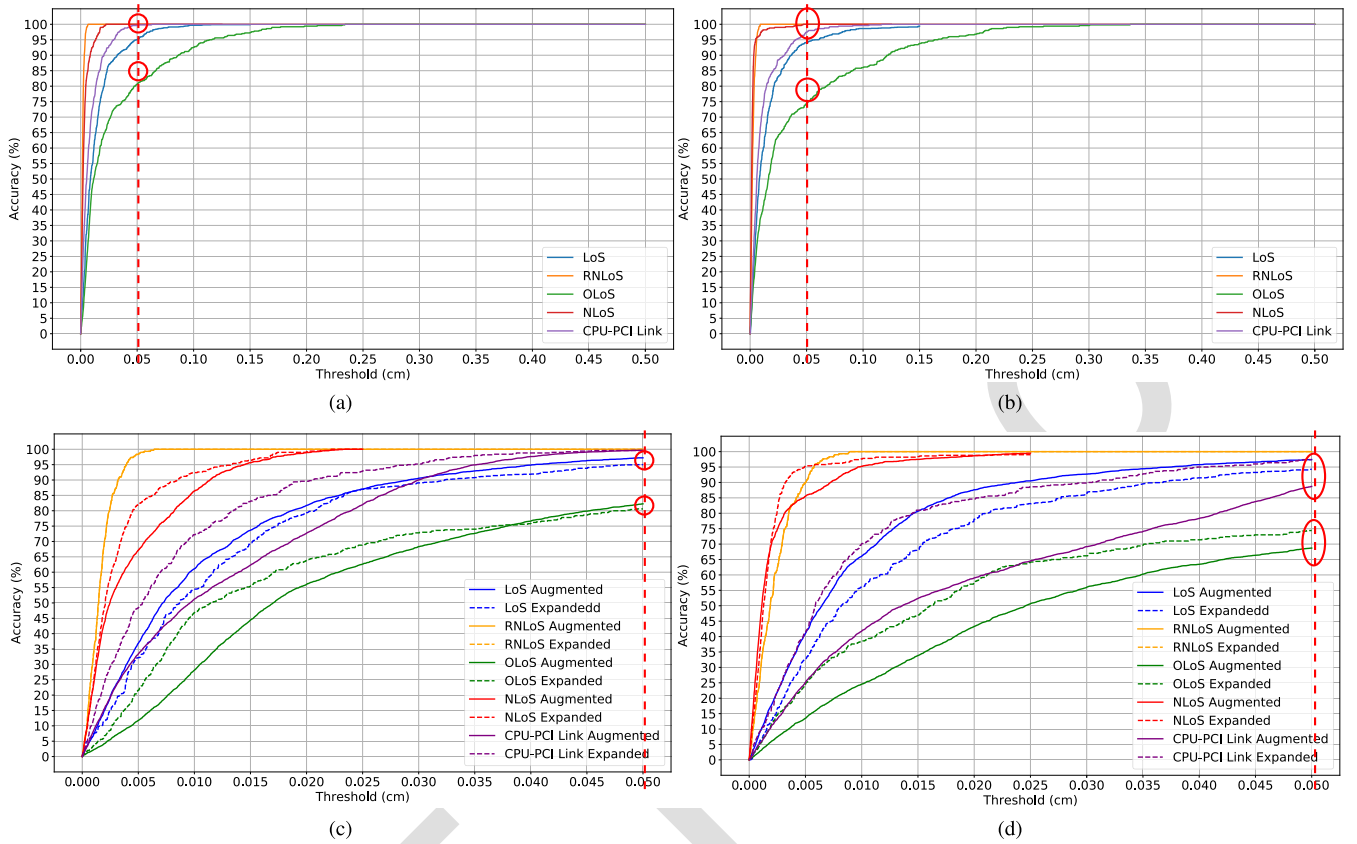


Fig. 8. Classification accuracy of (a) ResNet-based feature concatenated model tested on the expanded measured dataset, (b) MLP-based model tested on the expanded measured dataset, (c) ResNet-based feature concatenated model tested on the averaged interpolated dataset, and (d) MLP-based model tested on the averaged interpolated dataset for the attribute prediction under each potential scenario with respect to the selected threshold.

model used in comparison consists of a five-layer feature extraction block, a three-layer scenario prediction block, and another three-layer attribute prediction block, as shown in Fig. 9. Like our ResNet-based feature concatenated model, the MLP-based model also performs perfectly on the scenario classification task. Fig. 8(b) shows the classification accuracy of the ResNet-based feature concatenated model on the attribute prediction for each potential scenario with respect to the classification threshold on the expanded measured dataset. It can be seen from the plot that the MLP-based model also performs perfectly for the attribute prediction with nearly 100% achieved classification accuracy under all potential scenarios. However, by comparing Fig. 8(a) with (b), it can be observed that our ResNet-based feature concatenated model converges with tighter thresholds than the MLP-based model. Additionally, the classification accuracy achieved by the ResNet-based feature-concatenated model is higher than that of the MLP-based model on LoS, OLoS, and CPU-PCI link with a selected threshold of 0.05 cm, which indicates that the ResNet-based feature-concatenated model is more robust and performs better than the MLP-based model.

In addition to the evaluation on the expanded measured dataset, we also tested and compared our proposed ResNet-based feature concatenated model with the MLP-based model on the averaged interpolated dataset with higher resolution. Again, both models achieve 100% accuracy for scenario classification. However, while the ResNet-based

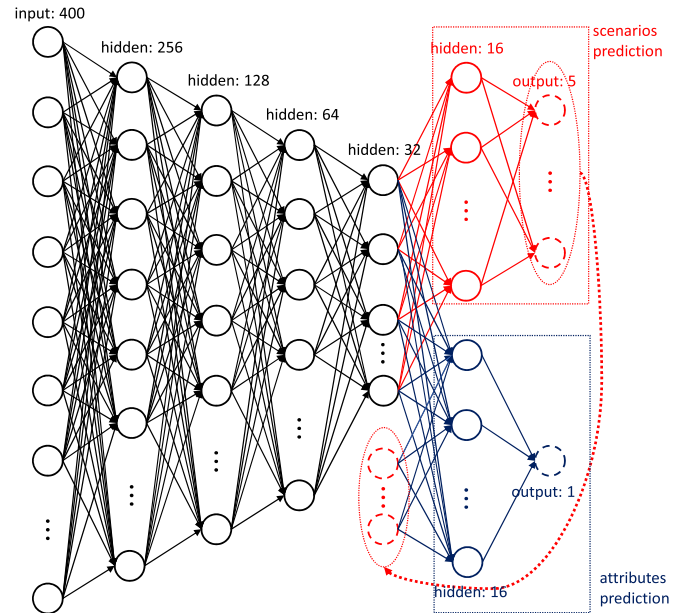


Fig. 9. Structure of the MLP-based network model.

feature concatenated model keeps constant performance for the prediction of the attribute under each scenario, the MLP-based model suffers a performance degradation on the averaged

interpolated dataset. Fig. 8(c) and (d) show the classification accuracy of our ResNet- and the MLP-based feature concatenated network models, respectively, for the attribute prediction under each potential scenarios with respect to the growing threshold on the averaged interpolated dataset and compare with the performances of the models on the measured expanded dataset. The selected threshold is in the range of zero to half of the higher-resolution step increment, which is 0.025 cm for NLoS and 0.05 cm for other scenarios. As shown in Fig. 8(c), it can be seen that the convergence rates of our proposed ResNet-based feature concatenated model on the expanded measured dataset (lower resolution) are very close to those on the averaged interpolated dataset (higher resolution), except for the CPU-PCI link which may because this scenario is much more complicated than others. With the selected threshold of 0.05 cm, as circled in red in these two plots, the classification accuracy of the ResNet base feature concatenated model tested on the averaged interpolated dataset is the same as that of the model tested on the expanded measured dataset. In contrast, it can be noticed from Fig. 8(d) that the convergence rates of the MLP-based model drop a lot on the augmented dataset for nearly all the scenarios, especially for the OLoS and CPU-PCI link. In addition, with the threshold of 0.05 cm, a 9% and a 5% drop of the classification accuracy for OLoS and CPU-PCI link, respectively, can be observed from the plots as circled in red. The performance degradation of the MLP-based model suggests that the range of the deviation between the MLP predicted and the correct value is larger than 0.05 for most attributes under the OLoS and CPU-PCI link, and the MLP-based model performs worse on the averaged interpolated data as the distance between interpolated neighbors are relatively short (half λ for NLoS and λ for others) and these neighbors are very similar to each other. The performance difference of our proposed ResNet- and MLP-based feature concatenated network models indicates that our proposed ResNet-based feature concatenated model is more robust and reliable on the dataset with higher resolution.

B. Model Analysis

To explain the performance difference of these two models and have a better understanding of the wireless channel and the deep learning models, FAM and Grad-CAM are applied on the MLP- and our proposed ResNet-based feature concatenated network models, respectively. To understand how the MLP-based network model makes predictions, we visualize a feature activation map by taking the absolute value of the average over the 256 neurons in the first linear layer as follows:

$$|\mathbf{w}_{\text{ave}}| = \left| \frac{1}{256} \sum_{i=1}^{256} \mathbf{w}_{i,*} \right| \quad (1)$$

where $\mathbf{W} \in R^{256 \times 400}$ represents the weights of the first linear layer [25]. FAM provides an illustration of which parts of the input are emphasized by an average neuron in the first layer to make decisions. As shown in Fig. 10, we highlighted a PDP measured under the OLoS scenario with the calculated $\mathbf{w}_{\text{ave}}$. It can be seen from the plot that the model focuses on the peaks with the excess delay of 0, 1.835, and 3.837 ns to make

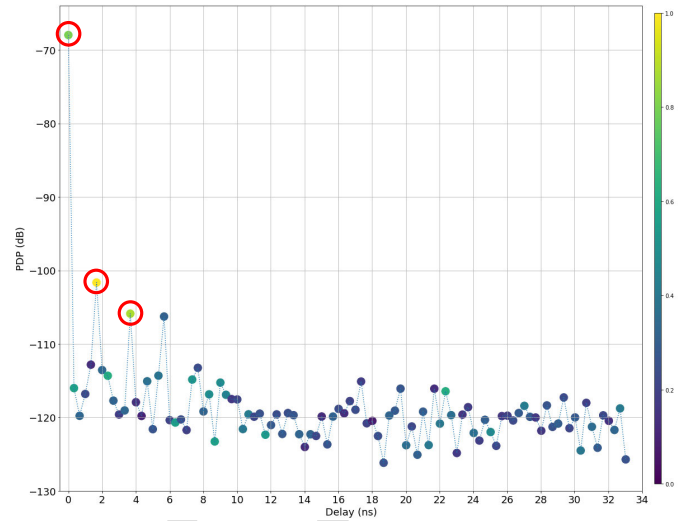


Fig. 10. Contributing regions of the PDP measured under the OLoS scenarios for the MLP-based network model where the attention of the neurons as measured by the FAM is denoted by the color of each point.

decisions, which are circled in red. Instead of the first peak (generated from the signal traveling directly from Tx to Rx) with the most power intensity, the model focuses more on the second and the third peaks (generated from the signal bouncing back and forth between the transceiver sidewalls [16]), which indicates that the model weights more on the multipath signals. This follows intuition since varying the distance between the parallel DIMMs under the OLoS scenario leaves the intensity of the direct traveling signal unchanged while the multipath signal intensity would fluctuate due to interactions of the DIMMs with the Fresnel zones [16]. In spite of this, the intensity of the first peak still influences the model decisions as it has strong differences when we are between different scenarios. Since the intensities of these peaks where the MLP model focuses are very different for the PDP under each scenario [16], the MLP-based model works perfectly on both the expanded measured and averaged interpolated datasets for the scenarios classification task.

The feature map also explains the performance degradation of the MLP-based model on the averaged interpolated dataset. Fig. 11(a) compares the PDP under OLoS scenario with the attributes from 5.2 to 5.6 cm and the part of the confusion matrix of the MLP-based model at these locations. It can be seen from the confusion matrix that the MLP-based model fails in these locations as the classification accuracy is less than 20%. Also, it can be seen that the PDP are very similar and the intensities of the peaks on which the model focuses, as circled in red in the figure, are very close to each other; which explains the failure of the model. As a counterexample, we also plot and compare the PDP with the attributes from 2.0 to 2.3 cm and the part of the confusion matrix at these locations as shown in Fig. 11(b). This figure shows that the MLP-based model performs much better at these locations as the classification accuracy is greater than 80%. It can be observed that the PDP is more easily differentiable between those differences at the peak locations where the model focuses. The contributing

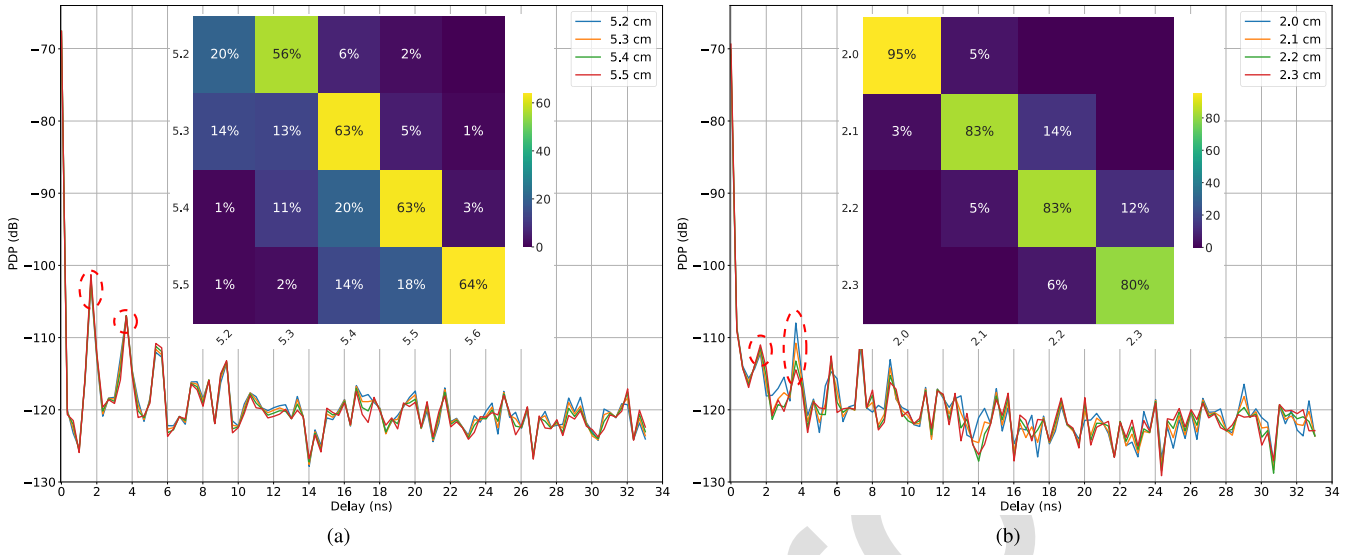


Fig. 11. PDP and confusion matrix of the MLP model under OLoS scenario with (a) attributes from 5.2 to 5.5 cm and (b) attribute from 2.0 to 2.3 cm.

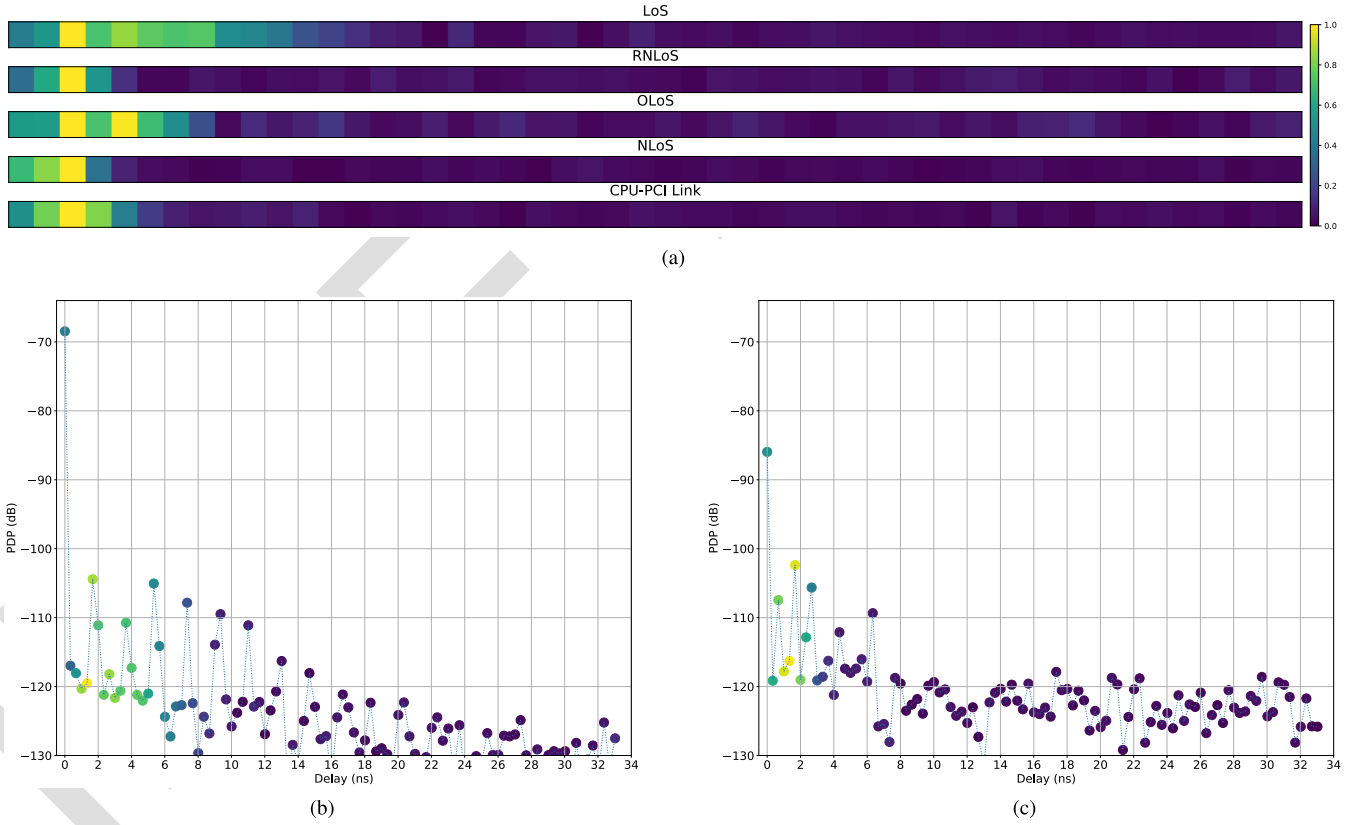


Fig. 12. (a) Heatmap of calculated $L^c_{Grad-CAM}$ with randomly selected PDP as the input for each scenario. (b) Contributing regions colored onto the randomly-selected PDP for the LoS channel. (c) Contributing regions of the randomly-selected PDP for the CPU-PCI link.

regions shown in Fig. 10 also apply to other potential scenarios as the weights of the neurons are fixed for the well trained MLP-based model.

Similarly, we also want to understand how the ResNet-based feature concatenated model makes predictions. However, the FAM is not suitable here as the ResNet-based feature concatenated model has a very different structure than the MLP-based model where convolutional layers instead of fully connected

layers are used to extract the features from the input data. Class activation mapping (CAM) was first proposed to produce a localization map for a classification CNN with a specific architecture, where global average pooled convolutional feature maps are fed directly to the output layer, by projecting back the weights of the output layer on to the convolutional feature maps [48]. A CAM can localize the class-distinguishing region of an input without positional supervision nor the requirement

of a backward pass. However, the constraint on the model architecture is restrictive so that it may not be useful for our case as the GAP layer is connected by another two fully connected networks (scenarios and attributes prediction block) in which the spatial information retained by previous convolutional layers is lost. The Grad-CAM was proposed to address the disadvantage of CAM as it uses the gradient information from the last convolutional layer of the CNN to assign importance values to each neuron for a particular class of interest [26]. The Grad-CAM is used here to provide visual explanations for the decisions made by our ResNet-based feature concatenated model. To derive the class-discriminative localization map $L_{Grad-CAM}^c$ for any class c , the gradient of the score for class c , y^c , with respect to feature map activations A^k of the last convolutional layer are first calculated, $(\partial y^c / \partial A^k)$. These gradients are then global average pooled over the last dimensions (indexed by i) to generate the neuron importance weights α_k^c as follows [26]:

$$\alpha_k^c = \frac{1}{Z} \sum_i \frac{\partial y^c}{\partial A_i^k} \quad (2)$$

where Z is the number of entries in the feature map ($Z = \sum_i 1$). The weight α_k^c captures the importance of feature map k for a target class c . The localization map $L_{Grad-CAM}^c$ is derived by combining weighted forward activation maps and taking the absolute value of it as follows:

$$L_{Grad-CAM}^c = \left| \sum_k \alpha_k^c A^k \right| \quad (3)$$

where K is the number of feature maps of the last convolutional layer. This result provides a coarse heatmap of the same size as the convolutional feature maps in the last convolutional layer. To highlight the regions where the model has more attention to make decisions, we first enlarged the localization map $L_{Grad-CAM}^c$ to the size of input PDP and highlighted the PDP with the enlarged $L_{Grad-CAM}^c$ for each scenario. Instead of using ReLU for the Grad-CAM localizations as described in [26], we take the absolute value as we are interested in the features that have any influence (not just positive influence) on the class of interest. It has been proven that Grad-CAM generalizes CAM for CNN-based architectures as the class features weights ω_k^c used in CAM are proportional to the neuron importance weights α_k^c used by Grad-CAM [26]. This generalization allows us to provide a visual explanation for our ResNet-based feature concatenated model which has complicated structures connected behind the GAP layer. Fig. 12(a) shows the heatmap of the calculated $L_{Grad-CAM}^c$ with randomly selected PDP as the input for each potential scenario. Fig. 12(b) and (c) highlight the contributing regions of the randomly selected PDP for LoS, and CPU-PCI link, respectively, based on the corresponding $L_{Grad-CAM}^c$. It can be observed from the plots that, similar to the MLP-based model, the ResNet-based feature concatenated model also focuses on the first few peaks of the PDP with specific emphasis depending on the scenario. Interestingly, both models pay more attention to the multipath signals than the direct traveling signal. However, besides the peaks,

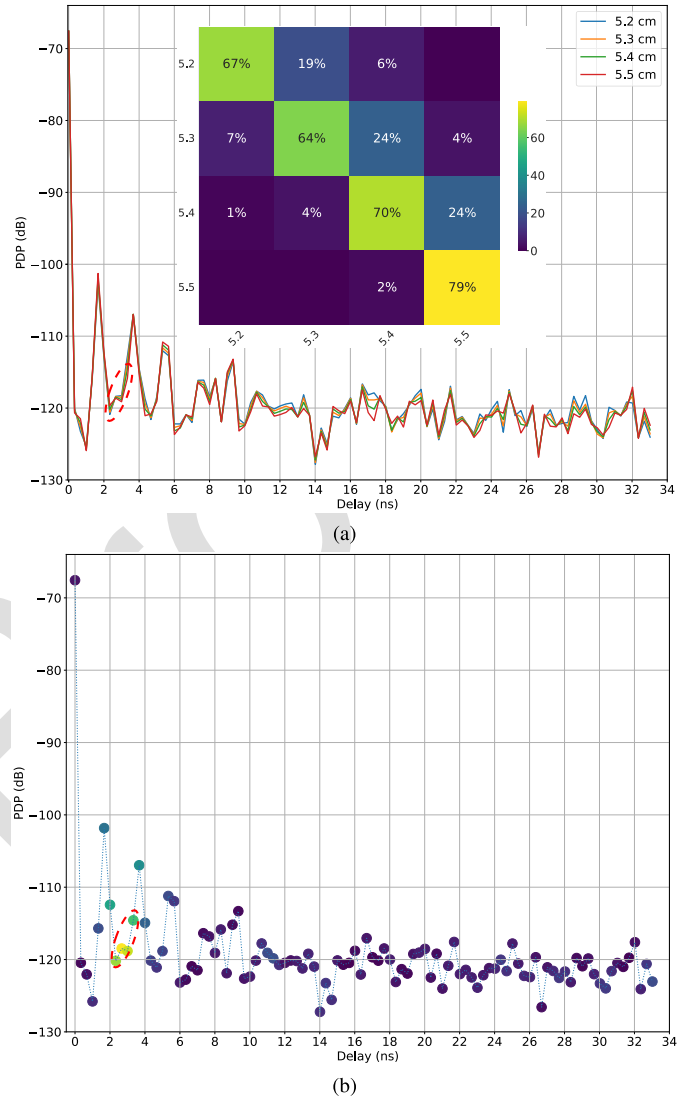


Fig. 13. (a) PDP and confusion matrix of the ResNet-based feature concatenated model under OLoS scenario with attributes from 5.2 to 5.5 cm. (b) Highlighted contributing regions of the PDP with the attribute of 5.3 cm under OLoS scenario.

the ResNet-based feature concatenated model also focuses on the valleys between those peaks even with stronger attention. This indicates that these valleys can also provide important information on the wireless channel and could also explain why the ResNet-based feature concatenated model performs more robustly than the MLP model. Fig. 13(a) shows the PDP and the confusion matrix of the ResNet-based feature concatenated model, again, under the OLoS scenario with the attribute ranging from 5.2 to 5.5 cm. Compared with the confusion matrix of the MLP model shown in Fig. 11(a), it can be concluded that the ResNet-based feature concatenated model performs much better than the MLP model at these locations as the classification accuracy of the ResNet-based feature concatenated model is greater than 64% while that of the MLP model is even less than 20%. Fig. 13(b) shows the contributing regions of the PDP with the attribute of 5.3 cm under the OLoS scenario. As shown in

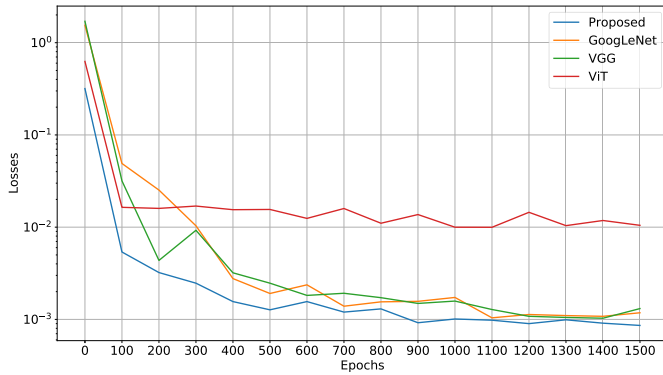


Fig. 14. Testing losses of our proposed ResNet-based feature concatenated model, GoogLeNet, VGG, and ViT on the same testing dataset throughout the training process.

this plot, besides the peaks due to the multipath signals, the model has even more attention on the valleys between those peaks where the PDP with the attributes from 5.2 to 5.5 cm can be clearly distinguished with each other as circled in red in Fig. 13(a), which explains why the ResNet-based feature concatenated model performs better at these locations.

Our proposed model performs better than other models as it has strong attention on the valleys between the multipath peaks. To show this, we compared the performance of our proposed ResNet-based feature concatenated model with some other mainstream deep learning models which include the GoogLeNet, visual geometry group (VGG), and vision transformer (ViT). Fig. 14 shows the testing losses of our proposed model, GoogLeNet, VGG, and ViT when evaluated on the same testing dataset throughout the training process. As shown in the dataset, our ResNet-based feature concatenated model outperforms other models, which illustrates the superiority of our proposed model.

V. CONCLUSION

This article presents the application of deep learning method on property prediction of the THz wireless channel in a motherboard desktop environment. Measurements were performed on a motherboard in a desktop size metal cavity. The PDP data which contains the channel information fed into the deep learning models was collected under five environment scenarios. For each scenario, measurements were taken several times with different attributes. To train the proposed deep learning model robustly and evaluate their performance comprehensively, data augmentation is applied on the measured PDP to expand the data size and to improve the resolution. A ResNet-based feature concatenated neural network model was proposed to perform the channel prediction. The generalized model structure consists of three blocks which are feature extraction, scenarios prediction, and attributes prediction block. The proposed model is evaluated and compared with an MLP-based model on both expanded measured and averaged interpolated datasets. Both models perform well on the expanded measured dataset with nearly perfect prediction

on both scenarios and attributes. Our proposed model outperforms the MLP-based model on the averaged interpolated dataset as the MLP model has a drop of classification accuracy on the attribute prediction task for each scenario with a fixed threshold and converges slower with respect to the larger thresholds, while our ResNet-based feature concatenated model keeps constant performance. The analyses are made with FAM and Grad-CAM applied on the MLP- and our ResNet-based feature concatenated models, respectively. It is shown from the FAM that the MLP model has a strong focus on the peaks of the PDP corresponding to multipath signals. However, some of the adjacent interpolated data are very close to each other in the regions on which the MLP focuses, which leads to performance degradation (or failure) on these areas. Also, as seen from the Grad-CAM, the ResNet-based feature concatenated model also has strong attention on the valleys between peaks where the neighboring interpolated data can be distinguished better, which accounts for its superior performance. Better performance in comparison with other mainstream deep learning models also proves the superiority of our proposed ResNet-based feature concatenated model.

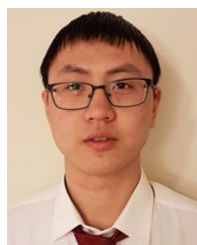
ACKNOWLEDGMENT

The views and findings in this paper are those of the authors and do not necessarily reflect the views of NSF.

REFERENCES

- [1] J. Fu, P. Juyal, and A. Zajic, "THz channel characterization of chip-to-chip communication in desktop size metal enclosure," *IEEE Trans. Antennas Propag.*, vol. 67, no. 12, pp. 7550–7560, Dec. 2019.
- [2] S. Kim and A. Zajic, "Characterization of 300-GHz wireless channel on a computer motherboard," *IEEE Trans. Antennas Propag.*, vol. 64, no. 12, pp. 5411–5423, Dec. 2016.
- [3] S. Kim and A. Zajic, "Statistical modeling and simulation of short-range device-to-device communication channels at sub-THz frequencies," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6423–6433, Sep. 2016.
- [4] J. Fu, P. Juyal, and A. Zajic, "Modeling of 300 GHz chip-to-chip wireless channels in metal enclosures," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3214–3227, May 2020.
- [5] J. Fu, P. Juyal, B. B. Yilmaz, and A. Zajic, "Investigation of resonance based propagation loss modeling for THz chip-to-chip wireless communications," in *Proc. 14th Eur. Conf. Antennas Propag. (EuCAP)*, 2020, pp. 1–5.
- [6] C.-L. Cheng, S. Sangodoyin, and A. Zajic, "THz cluster-based modeling and propagation characterization in a data center environment," *IEEE Access*, vol. 8, pp. 56544–56558, 2020.
- [7] G. P. Fettweis, N. Ul Hassan, L. Landau, and E. Fischer, "Wireless interconnect for board and chip level," in *Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE)*, 2013, pp. 958–963.
- [8] T. Kurner, A. Fricke, S. Rey, P. Le Bars, A. Mounir, and T. Kleine-Ostmann, "Measurements and modeling of basic propagation characteristics for intra-device communications at 60 GHz and 300 GHz," *J. Infr., Millim., Terahertz Waves*, vol. 36, no. 2, pp. 144–158, Feb. 2015.
- [9] A. Fricke, M. Achir, P. Le Bars, and T. Kurner, "A model for the reflection of terahertz signals from printed circuit board surfaces," *Int. J. Microw. Wireless Technol.*, vol. 10, no. 2, pp. 179–186, Mar. 2018.
- [10] J. Fu, "THz wireless channel characterization and modeling for chip-to-chip communication in computing systems," Ph.D. dissertation, Georgia Inst. Technol., Atlanta, GA, USA, 2022.
- [11] IT Association. *Infiniband Roadmap*. [Online]. Available: <https://www.infinibandta.org/infiniband-roadmap/>

- [12] C. Huang et al., "Artificial intelligence enabled radio propagation for communications—Part II: Scenario identification and channel modeling," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3955–3969, Jun. 2022.
- [13] D. He et al., "Stochastic channel modeling for kiosk applications in the terahertz band," *IEEE Trans. Terahertz Sci. Technol.*, vol. 7, no. 5, pp. 502–513, Sep. 2017.
- [14] R. Piesiewicz, C. Jansen, D. Mittleman, T. Kleine-Ostmann, M. Koch, and T. Kurner, "Scattering analysis for the modeling of THz communication systems," *IEEE Trans. Antennas Propag.*, vol. 55, no. 11, pp. 3002–3009, Nov. 2007.
- [15] S. Priebe, M. Jacob, C. Jansen, and T. Kurner, "Non-specular scattering modeling for THz propagation simulations," in *Proc. 5th Eur. Conf. Antennas Propag. (EUCAP)*, 2011, pp. 1–5.
- [16] J. Fu, P. Juyal, and A. Zajic, "Generalized modeling and propagation characterization of THz wireless links in computer desktop environment," *Radio Sci.*, vol. 57, no. 4, pp. 1–26, Apr. 2022.
- [17] J. Fu, P. Juyal, E. J. Jorgensen, and A. Zajic, "Comparison of statistical and deep learning path loss model for motherboard desktop environment," in *Proc. 16th Eur. Conf. Antennas Propag. (EuCAP)*, Mar. 2022, pp. 1–5.
- [18] C. Huang et al., "Artificial intelligence enabled radio propagation for communications—Part I: Channel characterization and antenna-channel optimization," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3939–3954, Jun. 2022.
- [19] M. Uccellari et al., "On the application of support vector machines to the prediction of propagation losses at 169 MHz for smart metering applications," *IET Microw., Antennas Propag.*, vol. 12, no. 3, pp. 302–312, Feb. 2018.
- [20] J. Salmi, A. Richter, and V. Koivunen, "Detection and tracking of MIMO propagation path parameters using state-space approach," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1538–1550, Apr. 2009.
- [21] Y.-T. Ma, K.-H. Liu, and Y.-N. Guo, "Artificial neural network modeling approach to power-line communication multi-path channel," in *Proc. Int. Conf. Neural Netw. Signal Process.*, Jun. 2008, pp. 229–232.
- [22] Q. Zheng et al., "Channel non-line-of-sight identification based on convolutional neural networks," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1500–1504, Sep. 2020.
- [23] M. I. AlHajri, N. T. Ali, and R. M. Shubair, "Classification of indoor environments for IoT applications: A machine learning approach," *IEEE Antennas Wireless Propag. Lett.*, vol. 17, no. 12, pp. 2164–2168, Dec. 2018.
- [24] M. Yang et al., "Machine-learning-based scenario identification using channel characteristics in intelligent vehicular communications," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 3961–3974, Jul. 2021.
- [25] E. J. Jorgensen, F. T. Werner, M. Prvulovic, and A. Zajic, "Deep learning classification of motherboard components by leveraging EM side-channel signals," *J. Hardw. Syst. Secur.*, vol. 5, no. 2, pp. 114–126, Jun. 2021.
- [26] R. R. Selyaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.
- [27] Virginia Diodes. *Custom Transmitters*. [Online]. Available: <https://vadiodes.com/en/products/custom-transmitters>
- [28] UI Communications. *Phase Locked Dielectric Resonator Oscillator*. [Online]. Available: <https://www.ultra.group/media/2081/ultra-herley-series-pdro.pdf>
- [29] Millimeter. *Frequency Multiplier—33 to 50 GHz*. [Online]. Available: <https://nordengroup.com/product-group/frequency-multiplier-33-to-50-ghz/>
- [30] Virginia Diodes. *Multipliers—VDI Model: WR6.5X3*. [Online]. Available: <https://www.vadiodes.com/en/wr6-5x3>
- [31] Virginia Diodes. *Mixers—VDI Model: WR2.8SHM*. [Online]. Available: <https://www.vadiodes.com/en/wr2-8shsm>
- [32] S. Kim and A. G. Zajic, "A path loss model for 300-GHz wireless channels," in *Proc. IEEE Antennas Propag. Soc. Int. Symp. (APSURSI)*, Jul. 2014, pp. 1175–1176.
- [33] M. Lin, Q. Chen, and S. Yan, "Network in network," 2013, *arXiv:1312.4400*.
- [34] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*.
- [35] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. 14th Int. Conf. Artif. Intell. Statist.*, 2011, pp. 315–323.
- [36] M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex," *Nature Neurosci.*, vol. 2, no. 11, pp. 1019–1025, Nov. 1999.
- [37] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [39] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, Jan. 1989.
- [40] M. Kalakh, N. Kandil, and N. Hakem, "Neural networks model of an UWB channel path loss in a mine environment," in *Proc. IEEE 75th Veh. Technol. Conf. (VTC Spring)*, May 2012, pp. 1–5.
- [41] N. Zaarour, N. Kandil, N. Hakem, and C. Despins, "Comparative experimental study on modeling the path loss of an UWB channel in a mine environment using MLP and RBF neural networks," in *Proc. Int. Conf. Wireless Commun. Underground Confined Areas*, Aug. 2012, pp. 1–6.
- [42] N. Zaarour, N. Kandil, and N. Hakem, "An accurate neural network approach in modeling an UWB channel in an underground mine," in *Proc. IEEE Antennas Propag. Soc. Int. Symp. (APSURSI)*, Jul. 2013, pp. 1608–1609.
- [43] N. Zaarour, S. Affes, N. Kandil, and N. Hakem, "Comparative study on a 60 GHz path loss channel modeling in a mine environment using neural networks," in *Proc. IEEE Int. Conf. Ubiquitous Wireless Broadband (ICUWB)*, Oct. 2015, pp. 1–4.
- [44] W. Hu, S. Geng, and X. Zhao, "Mm-wave 60 GHz channel fading effects analysis based on RBF neural network," in *Proc. 5th Int. Conf. Comput. Commun. Syst. (ICCCS)*, May 2020, pp. 589–593.
- [45] A. Neskovic, N. Neskovic, and D. Paunovic, "Macrocell electric field strength prediction model based upon artificial neural networks," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1170–1177, Aug. 2002.
- [46] J. A. Romo, I. F. Anitzine, and F. P. Fontan, "Application of neural networks to field strength prediction for indoor environments," in *Proc. 1st Eur. Conf. Antennas Propag.*, Nov. 2006, pp. 1–6.
- [47] H. Zhou, F. Wang, and C. Yang, "Application of artificial neural networks to the prediction of field strength in indoor environment for wireless LAN," in *Proc. Int. Conf. Wireless Commun., Netw. Mobile Comput.*, 2005, pp. 1189–1192.
- [48] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2921–2929.



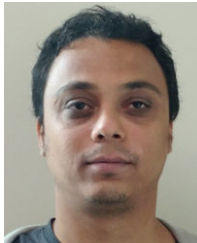
Jinbang Fu was born in Nanchang, China. He received the B.Sc. degree (summa cum laude) in electrical and computer engineering from the University of Massachusetts, Amherst, MA, USA, in 2017. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the Georgia Institute of Technology, Atlanta, GA, USA.

He is a Graduate Research Assistant with the Electromagnetic Measurements in Communications and Computing Laboratory, Georgia Institute of Technology. His current research interests include THz wireless channel measurements and modeling and THz wireless communication system design for chip-to-chip communication.



Erik J. Jorgensen (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Wisconsin, Madison, WI, USA, in 2017, and the M.Sc. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2019, where he is currently pursuing the Ph.D. degree in machine learning with the Electromagnetic Measurements in Communications and Computing (EMC2) Laboratory.

His current research interests include explainable machine learning, deep learning methods, computer vision, and optimization.



Prateek Juyal received the B.E. degree in electronics and communication engineering from Kumaun University, Nainital, India, in 2007, the M.Tech. degree in digital communication from Guru Gobind Singh Indraprastha University, New Delhi, India, in 2009, and the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Winnipeg, MB, Canada, in 2017.

Currently, he is working as a Post-Doctoral Fellow with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. His current research interests include sensors, security, and wireless communication.



Alenka Zajić (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees from the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia, in 2001 and 2003, respectively, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2008.

Currently, she is an Associate Professor at the School of Electrical and Computer Engineering, Georgia Institute of Technology. Prior to that, she was a Visiting Faculty Member at the School of Computer Science, Georgia Institute of Technology, a Post-Doctoral Fellow at the Naval Research Laboratory, Washington, DC, USA, and a Design Engineer at Skyworks Solutions Inc, Irvine, CA, USA. Her research interests span areas of electromagnetic, wireless communications, signal processing, and computer engineering.

Dr. Zajić was a recipient of the 2017 NSF CAREER award, the Best Paper Award at MICRO 2016, 2012 Neal Shepherd Memorial Best Propagation Paper Award, the Best Student Paper Award at the IEEE International Conference on Communications and Electronics 2014, the Best Paper Award at the International Conference on Telecommunications 2008, the Best Student Paper Award at the 2007 Wireless Communications and Networking Conference, and the Dan Noble Fellowship in 2004, which was awarded by Motorola Inc. and the IEEE Vehicular Technology Society for quality impact on the area of vehicular technology.

AUTHOR QUERIES

PLEASE NOTE: We cannot accept new source files as corrections for your article. If possible, please annotate the PDF proof we have sent you with your corrections and upload it via the Author Gateway. Alternatively, you may send us your corrections in list format. You may also upload revised graphics via the Author Gateway.

Carefully check the page proofs (and coordinate with all authors); additional changes or updates **WILL NOT** be accepted after the article is published online/print in its final form. Please check author names and affiliations, funding, as well as the overall article for any errors prior to sending in your author proof corrections.

AQ:1 = The abbreviation FAM has been used for the terms “feature activation map” and “feature activation mapping” Please confirm which has to be followed.

AQ:2 = Please confirm or add details for any funding or financial support for the research of this article.

AQ:3 = Please provide the full current affiliation details (department name, name of university/institution, city, country, and postal code) for all the authors.

AQ:4 = The in-text citations of “Fig. 2” is out of order. Fig. 1 should be cited before Fig. 2. Please update the in-text citations.

AQ:5 = Please confirm the retention of the sentence in the acknowledgment section is correct.

AQ:6 = Please provide the accessed date for Refs. [11], [27], [28], [29], [30], and [31].

Deep Learning-Based THz Wireless Channel Property Prediction in Motherboard Desktop Environment

Jinbang Fu^{ID}, Erik J. Jorgensen^{ID}, *Graduate Student Member, IEEE*, Prateek Juyal^{ID},
and Alenka Zajić^{ID}, *Senior Member, IEEE*

Abstract—This article proposes a residual network (ResNet)-based feature concatenated neural network model to predict the type of scenario the channel is under and the attribute of the predicted scenario with power delay profile (PDP) as the inputs. The generalized model structure consists of three blocks for feature extraction, scenario prediction, and attribute prediction, respectively. The PDP data is collected from a motherboard desktop environment under five different physical arrangement scenarios. Within each scenario, data is collected several times while varying a different physical attributes for each scenario. Two steps of data augmentation are applied to expand the size and to improve the resolution (difference between the neighboring attributes) of the measured dataset for the robust training and thorough evaluation of the proposed model. The proposed model is evaluated and compared with a multilayer perceptron (MLP)-based model on an expanded measured and averaged interpolated dataset. It is shown that both models perform very well on the expanded measured dataset with nearly 100% prediction accuracy on both scenarios and attributes. The MLP-based model suffers performance degradation on the averaged interpolated dataset with up to a 9% drop of classification accuracy on attribute prediction tasks, while our ResNet-based feature concatenated model performs equally in both scenarios. Feature activation map (FAM) and grad-class activation mapping (Grad-CAM) approaches are applied to provide visual explanations highlighting characteristics of the input PDP used for model decisions. FAM shows that the MLP-based model focuses on the multipath generated peaks of the PDP where some interpolated neighboring data points cannot be distinguished. The Grad-CAM shows that the proposed ResNet-based feature concatenated model performs better because it has strong attention not only on the multipath peaks but also on the valleys between those peaks which hold distinguishing information.

Index Terms—Channel characterization, channel prediction, channel sounding, chip-to-chip wireless channels, THz communications.

I. INTRODUCTION

THE THz wireless link is considered as a replacement for wired connections in inter-chip communication for future computing systems [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]. A wireless link system alleviates the needs for

cable management and reduces the complexity of the computing system design [1], [2]. Operating at THz frequencies with its larger available bandwidth allows higher speed and lower latency communications. With smaller antenna sizes and reduced spacing at THz frequencies, spatial multiplexing can be applied to improve data rates to Terabit-per-second which is close to the data rate of current wired systems [2], [11]. To realize THz wireless inter-chip communication, comprehensive understanding, and identification of the corresponding wireless channels is required. Wireless channels are usually determined by the signal propagation environment and interactions with the objects in the environment, also known as the channel properties [12]. These channel properties define the characteristics of the wireless channel and affect the quality of possible wireless links. The prediction of channel properties is important for establishing inter-chip communication and allows for further adjustments to improve the communication link. With a precise prediction of channel properties, the appropriate transmission mode can be selected to adapt to the environment, which can help to provide more reliable communication with better performance [12].

To realize THz wireless inter-chip communication, channel modeling tools are needed and several THz wireless channel models have been proposed [8], [9], [13], [14], [15]. Based on the geometrical distribution of the environment, statistical channel models have been proposed for short-range device-to-device channels and THz wireless inter-chip communication channels [3], [4]. Constant rate models have been applied on THz wireless channels in a datacenter and in the motherboard desktop environment by modifying the classic Saleh–Valenzuela (S-V) model [6], [16]. These classic analytical techniques can provide good estimations of the wireless channel with simple math equations. However, these techniques are generally not suitable for the prediction and identification of channel properties because the simplifications introduced in the models may not capture more complex multipath environment properties [17]. Moreover, the large number of potential links, obstructions, scattering, and so on for the inter-chip communication in motherboard desktop environment requires a large amount of data for a thorough channel characterization. These classical techniques are less suitable for the analysis of large datasets as they might miss important relations within the data [18].

Manuscript received 16 June 2022; revised 20 March 2023; accepted 14 May 2023. This work was supported in part by the NSF under Grant 1651273. (Corresponding author: Jinbang Fu.)

The authors are with ??? (e-mail: jfu72@gatech.edu).
Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TAP.2023.3278831>.

Digital Object Identifier 10.1109/TAP.2023.3278831

In contrast, machine learning (ML) is a branch of artificial intelligence that learns a model from data to make decisions and has achieved breakthroughs in application areas such as image processing, natural language processing, and data mining [18]. Due to their rise in popularity, ML techniques have been gradually applied to study the wireless propagation channel, such as channel characterization and channel modeling [10], [12]. Researchers have applied ML techniques on the estimation of propagation loss [17], [19], tracking multipath components [20], and derivation of channel statistical parameters [21]. Also, ML techniques are applied on channel property prediction as they have been shown to provide good performance in identifying small differences between measurements to characterize different properties. Several deep learning-based methods have been proposed for a nonline of sight (NLoS) recognition [22], indoor environments classification [23], and vehicle communication scenarios identification [24]. Considering the complicated propagation channel on the densely populated motherboard in a metal casing, where some hidden patterns due to scattering, reflections, and resonant cavity effects are difficult to identify by traditional techniques, ML methods are more likely to succeed for channel prediction in such an environment.

To establish a stable wireless link for THz wireless inter-chip communication in computing systems, it is necessary to understand the channel properties in advance. These properties include potential physical scenarios of the channel such as the link between CPU and PCI, the link between CPU and dual in-line memory module (DIMM), and so on, and certain attributes under those scenarios such as the height differences between the chips or the distances between the CPU and the PCI components. These channel properties could be predetermined by an ML model with the information given by the pilot signal.

The main contributions of this article are summarized as follows.

- 1) We apply deep learning methods to predict channel properties for THz chip-to-chip wireless communication in motherboard desktop environment. With PDP as the input, we propose a residual network (ResNet)-based feature concatenated model to predict the channel's physical scenario and an additional attribute of that predicted scenario.
- 2) The proposed model is evaluated and compared with a multilayer perceptron (MLP)-based model on the expanded measured dataset and the averaged interpolated dataset. Results show that our model performs very well on the expanded measured dataset with nearly perfect prediction of both scenarios and attributes. These results exemplify the validity of our proposed model and the feasibility of applying deep learning model for THz channel prediction on a densely-populated motherboard in a practical computing system. We show that our ResNet-based feature concatenated model performs more robustly than the MLP model as it has higher classification accuracy when predicting a specific scenario's attribute. Our ResNet-based feature concatenated model also has constant performance on both measured

and augmented dataset while the MLP model shows performance degradation on the averaged interpolated dataset.

- 3) feature activation mapping (FAM) and Gradient-weighted class activation mapping (Grad-CAM) are applied on the MLP and our ResNet-based feature concatenated models, respectively, to explain their performance differences by determining the contributing regions of the raw data that the models use to make their predictions [25], [26]. The results provide us a deeper understanding of the THz wireless channel in the motherboard desktop environment. We find that the MLP model only focuses on the peaks of the input PDP with the delay of 0, 1.835, 3.837 ns corresponding to multipath propagations, while the ResNet-based feature concatenated model has strong attention not only to those peaks but also on the valleys between those peaks. This implies that the valleys between the peaks generated by the multipath signals also contain important information related to the wireless channel that help distinguish different channel properties.

The rest of the article is organized as follows. Section II presents the channel measurements and the augmentation of the collected data. Section III illustrates the design of the ResNet-based feature concatenated neural network model. Section IV presents the analysis and model verification. Finally, Section V makes concluding remarks.

II. DATA MEASUREMENTS AND AUGMENTATIONS

In this section, we provide a brief discussion about THz wireless channel measurements in the motherboard desktop environment under five different potential scenarios: LoS, reflected non-LoS (RNLoS), obstructed-LoS (OLoS), non-LoS (NLoS), and a practical CPU-PCI link. Under each of these scenarios, measurements are performed several times while varying a different attribute for each scenario. Due to time and physical setup limitations, we also apply multiple data augmentation techniques on the measured data for training, validation, and analysis of our proposed models.

A. THz Channel Measurements

The measurement setup includes a Keysight N5224A PNA vector network analyzer (VNA), a Virginia diodes, Inc. (VDI) transmitter (Tx210), and a VDI receiver (Rx148) [27]. The VNA generates the input signal in the range of 10 MHz–12 GHz with the power level (P_{in}) of 0 dBm and feeds it into the VDI Tx210. In the VDI Tx210 transmitter, a 25 GHz signal is generated by a Herley-CTI phase-locked dielectric resonator oscillator (DPRO with 100 MHz reference crystal oscillator) [28]. It is amplified and frequency doubled by Norden N08-1975 [29] and then tripled by VDI WR6.5X3 [30]. This signal is then fed into a sub-harmonic mixer (WR2.8SHM) [31] which doubles the carrier frequency and mixes it with the VNA-generated baseband signal. This THz-range signal is then transmitted by a horn antenna in the range of 288–312 GHz. The same components are used to down-convert the signal, the difference is that the DPRO is

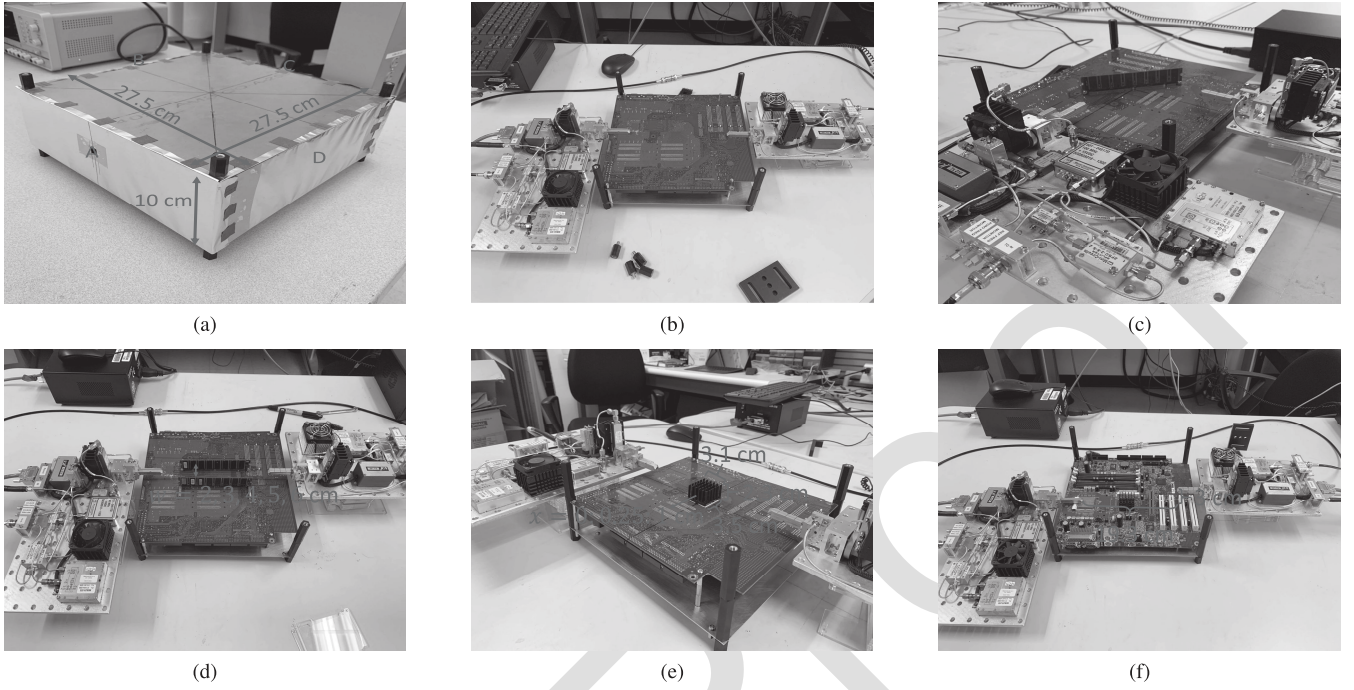


Fig. 1. Setups for the measurements of THz wireless channel in motherboard desktop environment. (a) Fabricated metal casing that houses each of (b) LoS, (c) RNLoS, (d) OLoS, (e) NLoS, and (f) CPU-PCI link measurement scenarios.

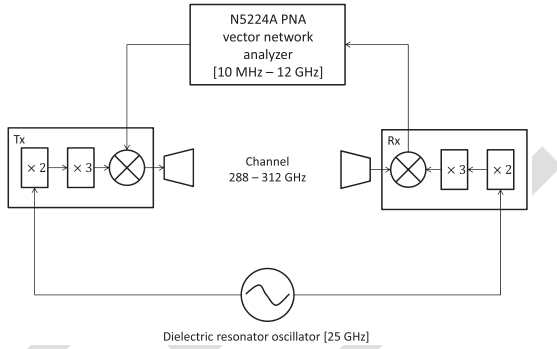


Fig. 2. Block diagram of the measurement system.

TABLE I
MEASUREMENT PARAMETERS

Parameter	Symbol	Value
Measurement points	N	801
Intermediate frequency bandwidth	Δf_{IF}	20 kHz
Average noise floor	P_N	-90 dBm
Input signal power	P_{in}	0 dBm
Start frequency	f_{start}	10 MHz
Stop frequency	f_{stop}	12 GHz
Bandwidth	B	11.99 GHz
Time domain resolution	Δt	0.083 ns
Maximum excess delay	τ_m	33.4 ns

tuned to 24.2 GHz. The received signal is then down-converted to an intermediate frequency (IF) of 9.6 GHz. The VNA then samples the upper sideband of the down-converted signal in the range of 9.61-21.6 GHz with 801 points. The block diagram of the measurement system is shown in Fig. 2 [1], [32] and all measurement parameters are summarized in Table I. Pyramid horn antennas with a gain that varies in the range of 22–23 dBi over the observed frequency band from 300 to 312 GHz were used in the measurements. The theoretical half-power beamwidth (HPBW) of the horn is about 12° in azimuth and elevation. The physical dimension of the horn aperture is 8.91 mm, which limits the far-field boundary to 15.88 cm at 300 GHz according to the Fraunhofer distance.

To measure the THz wireless channel in a motherboard desktop environment, we fabricated an aluminum metal cavity with the size of $27.5 \times 27.5 \times 10$ cm, which approximates the size of a desktop, with two square aluminum plates as the top and bottom walls [5]. The other four sides are

wrapped by aluminum foil and labeled as A–D, as shown in Fig. 1(a). As shown in Fig. 1(b), a motherboard was supported 4.2 cm over the bottom wall by brass hex standoffs inside the fabricated cases. For all channel environment scenarios, measurements are performed on the motherboard inside the metal cavity.

For the measurements of LoS links, as shown in Fig. 1(b), antennas of the transmitter (Tx) and receiver (Rx) were aligned with each other and inserted into the cavity through small holes drilled on the transceiver sidewalls (A and C) of the cavity. Measurements were performed on the backside of the motherboard to separate the influence of the other components. As the heights of components on the motherboard may differ and to investigate the ground effect of the motherboard, both

transceivers were elevated simultaneously in height during the measurements. To do this, the height of transceivers h was varied from 0 to 4.5 cm above the motherboard in increments of 0.3 cm while the distance between the transceivers was fixed at 19.3 cm.

As shown in Fig. 1(c), a DIMM was placed standing up on the center of the motherboard and diagonally aligned with transceivers which were positioned orthogonal to each other for the RNLoS measurements. The distance from the DIMM to each of the transceiver is 9.55 cm and the heights for both transceivers were fixed at 1.8 cm above the motherboard. Both flat and component sides of the DIMM were investigated during the measurements as both sides could possibly be used to establish the wireless link on the motherboard. For the OLoS link, measurements were performed to investigate the effects of parallel-plate structures on the motherboard. As shown in Fig. 1(d), two DIMMs were put in parallel on the center of the motherboard. when the distance between parallel plates, ω , is small, e.g., $\omega = 2$ or 3 cm, some parts of the parallel plates may be in the second or third region of the Fresnel ellipsoid space whose radius is 1.17 or 1.44 cm. The interceptions may lead to destructive or constructive effects due to the phase shift of the signal reflecting on the parallel plates [2]. Measurements were performed by varying the distances between the DIMMs from 2 to 6 cm with the step increment of 1 cm to simulate the variable distances between parallel structures on the motherboard. The heights of both transceivers were kept at 1.8 cm during the measurements.

The NLoS link was measured with a heatsink of size $3.5 \times 3.1 \times 1.3$ cm as an obstacle, as shown in Fig. 1(e). Between measurements, the heatsink was horizontally shifted away from the transceiver toward side D of the cavity, with a step increment of 0.25 cm to understand how the wireless link is affected by the gaps between the fins of the heatsink. The distance from the heatsink to each of the transceiver is 8.1 cm and the heights of both transceivers were kept at 1.8 cm.

Besides the investigations of these primary controlled scenarios, a practical CPU-PCI link was also measured on the densely populated motherboard. As shown in Fig. 1(f), Tx was mounted on top of the CPU and Rx was hung over the PCI slot with a 3 cm T-R vertical misalignment, which approximates the height difference between the CPU and the chip on the peripheral components. During the measurements, the horizontal T-R separations were gradually increased from 19.3 to 23.3 cm with the step increment of 1 cm as there are several PCI slots on the motherboard with different distances to the CPU. All the scenarios and corresponding variable attributes are summarized in Table II.

B. Data Augmentations

To robustly train our deep neural network model to precisely predict channel properties, we require a large dataset to feed into the model. However, measurements inside the metal cavity are very time-consuming as we need to rebuild the cavity every time we shift the transceivers or components for the measured attribute of each scenario. Additionally, due to physical constraints, we choose a relatively large step increment (low resolution) for the measurements under each

TABLE II
MEASURED POTENTIAL SCENARIOS AND CORRESPONDING ATTRIBUTES

LoS	The heights of transceivers
RNLoS	Flat/component side of the DIMM
OLoS	The distance between parallel DIMMs
NLoS	The movement of the heatsink
CPU-PCI link	The distance between the transceivers

scenario to ensure we can physically capture the data precisely. For instance, the heights of transceivers were varied during the LoS measurements by stacking plastic plates with a thickness of 0.3 cm under the transceivers. This means that information between the measured points is unknown to us. To validate the generalization ability of our proposed model, we also test them on a dataset with higher resolution. To that end, we apply data augmentation to expand the size of our dataset and increase its resolution by creating synthetic interpolated data from measured results. The tweaks of training examples in data augmentation make the model less likely to overfit on certain patterns, which improves the generalization ability of the model.

The power delay profile (PDP), which provides the intensity of a signal received through a multipath channel with respect to time delay, is fed into our deep neural network model for training and prediction. Measured PDP can be derived by squaring the inverse discrete Fourier transform (IDFT) of the measured frequency response, which itself is calculated by subtracting the S21 from the sum of antenna gains and the thru loss. Before any S21 measurements, the thru loss was first measured on the transceiver devices ten times. Then, as mentioned in II-A, we measured the S21 ten times at different attributes under each potential scenario. By combining all pairs of the ten thru loss and ten S21 measurements, we generate 100 samples for each attribute, as shown in Fig. 3(a). The expanded measured dataset was then separated into three groups by randomly separating 50%, 20%, and 30% of the data at each scenario and attribute to training, validation, and testing sets, respectively. Then the training and validation datasets were interpolated to a higher resolution for the scenarios which are measured in increments (all except the RNLoS scenario). For this augmentation, we applied linear, spline (spline of order 1), quadratic, cubic, PCHIP (piecewise cubic Hermite interpolating polynomial), and Akima interpolations to increase resolution down to steps of 0.1 cm (1λ) for LoS, OLoS, and CPU-PCI link measurements and to 0.05 cm (half λ) for the NLoS measurement. Interpolations were not applied on the data of RNLoS link because it is a binary dataset showing the difference between the flat or component side of the DIMM as the reflecting surface. Finally, we generated the average augmentation dataset by averaging the six interpolations for each attribute since we do not necessarily know which interpolation technique would be represent the unmeasured data. For example, data was measured at $h = 0, 0.3, 0.6, \dots, 4.5$ cm for the LoS propagation. Measurements were performed ten times at each height.

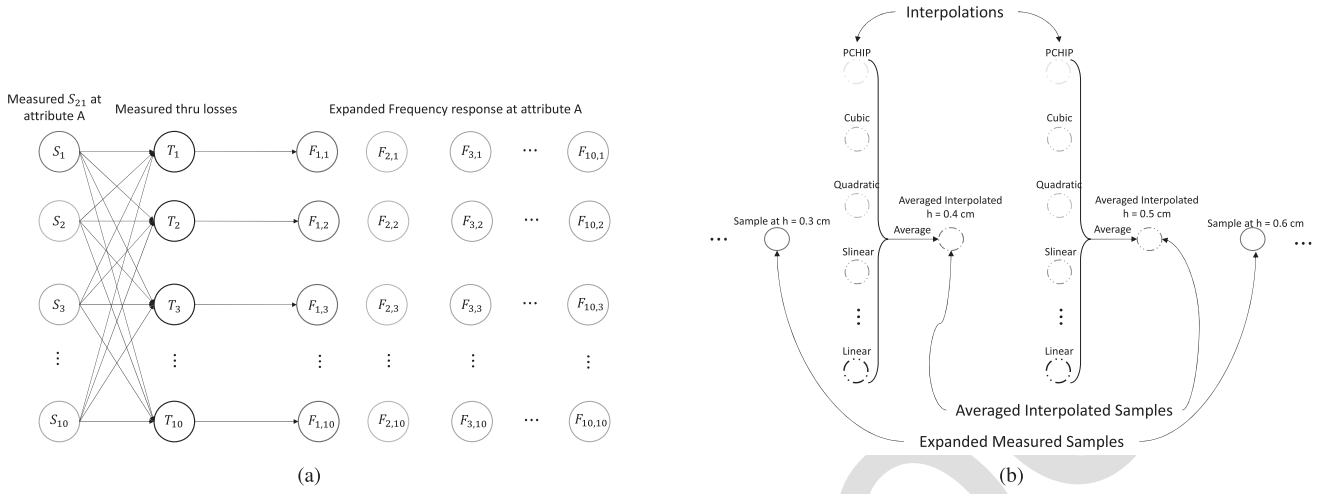


Fig. 3. Illustration of data augmentation: (a) data expansion and (b) averaged interpolation for data of LoS propagation.

With the data expansion, we generated 100 PDP samples at each height. As shown in Fig. 3(b), six different interpolation techniques were then applied to the training and validation datasets to improve the resolution from 0.3 to 0.1 cm. We then averaged those interpolations at each height. Doing so resulted in 50 and 30 PDP at each height $h = 0, 0.1, \dots, 4.5$ cm for the respective training and validation datasets, and 20 PDP at each of the same heights for the testing dataset. This averaged augmentation provides a single dataset (only one average PDP instead of six at each interpolated attribute) with higher resolution.

III. DEEP LEARNING MODEL

In this section, we introduce a ResNet-based feature concatenated neural network which is designed to predict properties of the THz wireless channel in the motherboard desktop environment. The general structure of the model is shown in Fig. 4. As shown in the figure, the proposed neural network model consists of three main blocks. The first is the “feature extraction” block which extracts the generalized features necessary to distinguish between scenarios and specific attributes of each scenario from the input PDP data. Those extracted features are then sent to the “scenarios prediction” block to predict the type of scenario the channel is under (LoS, RNLoS, OLoS, NLoS, and CPU-PCI link). The overall network design is called “feature concatenated” because the model’s scenario prediction output is then concatenated with the extracted features and fed to the “attributes prediction” block to estimate the attribute of that predicted scenario.

Our novel ResNet-based feature concatenated neural network model for the prediction of THz channel property is shown in Fig. 5. As shown in the figure, the feature extraction block of the proposed ResNet-based feature concatenated network model contains five sub-blocks, which are **b1**, **b2**, **b3**, **b4**, and a global average pooling (GAP) sub-block [33]. **b1** consists of a 64-channel 1×7 convolutional layer with the padding of 3 and a stride of 2, a batch normalization (BN) layer, rectified linear unit (ReLU) activation layer, and a 1×3 maximum pooling layer with a stride of 2. During

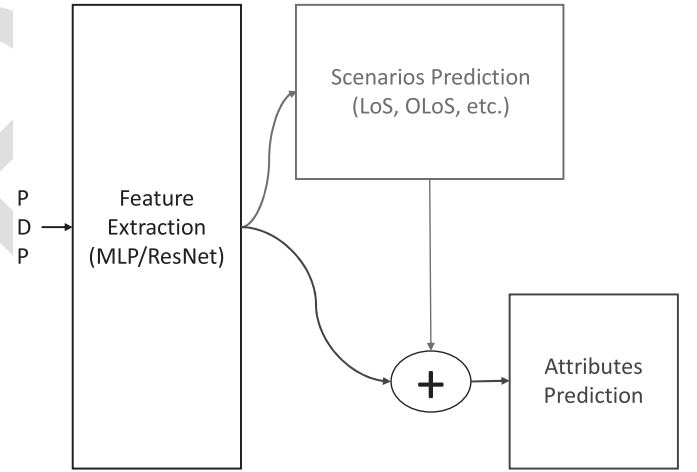


Fig. 4. Generalized model structure consisting of three main blocks.

training, the weights of intermediate convolution filters (e.g., the convolutional layers) may have strong variation between different layers, nodes in the same layer, and over time due to the updating model parameters. BN [34] is implemented to maintain stable optimization and reduce these strong variations that may hinder the convergence of the network. The ReLU [35] activation function is applied to increase the nonlinear expressiveness of the network, which simply outputs its input for positive values and 0 for negative values. A max-pooling layer is used to mitigate sensitivity of the network to noisy inputs and to spatially downsample representations [36]. **b1** increases the number of channels of the inputs from 1 to 64 while reducing the size of the data from 400 to 99. **b2–b4** are three ResNet building blocks shared with the same structures. Each of them consists of a residual block with an additional 1×1 convolutional layer to transform the inputs to the desired shape for the addition operation, and a residual block without the 1×1 convolutional layer. A residual block is the core element of a ResNet model, which guarantees the strictly increasing expressiveness of the network. The structure

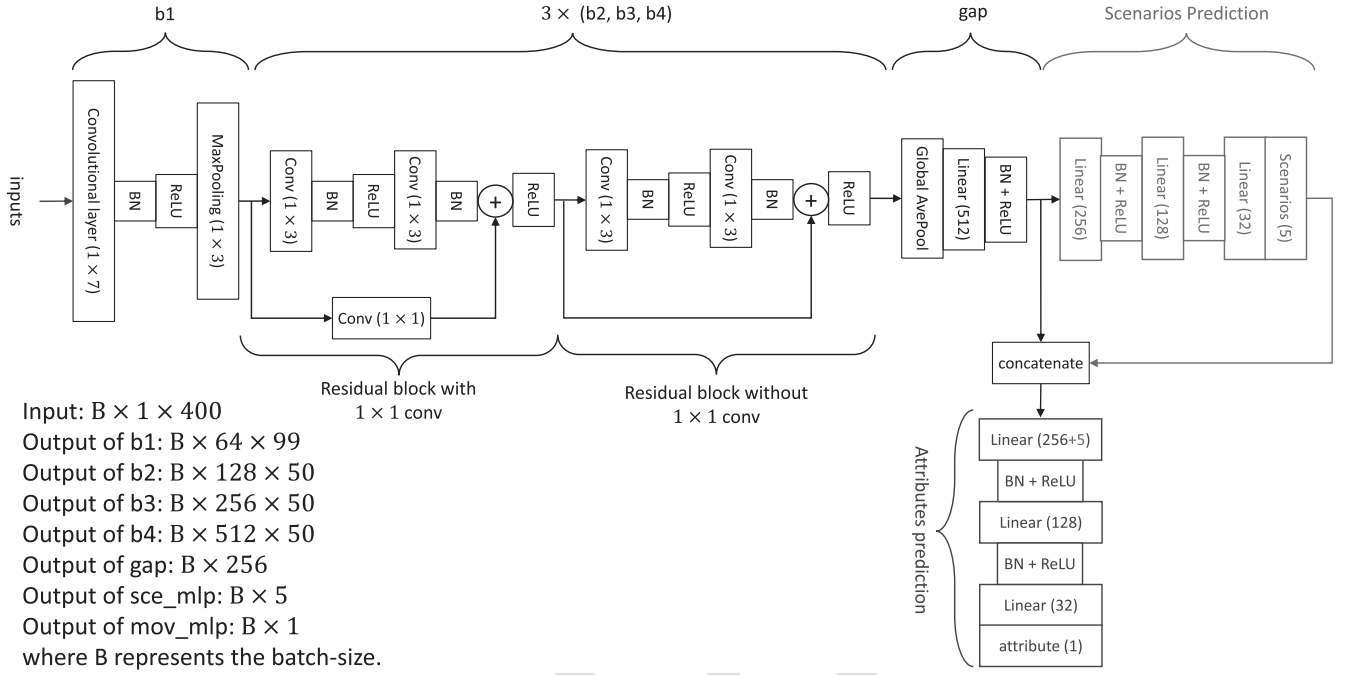


Fig. 5. Structure of the proposed ResNet-based feature concatenated network model.

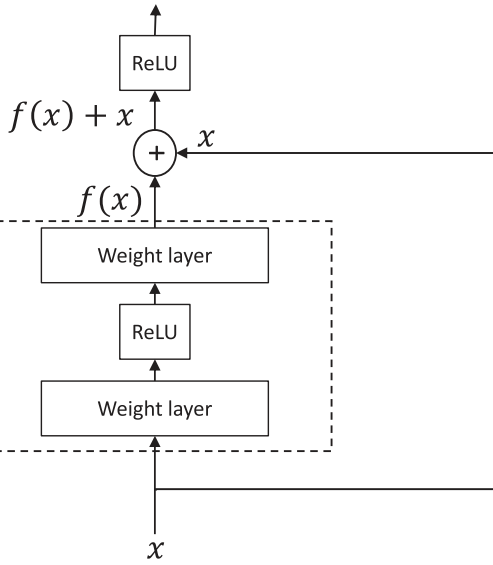


Fig. 6. Residual block.

than its shallower counterpart [37]. **b2** doubles the number of input channels and halves the size of input data. **b3** and **b4** keep doubling the number of input channels while maintaining the size of the input data. The GAP sub-block is composed of a GAP layer, which computes the average value across the entire matrix for each input channel, and a linear layer is connected behind, which halves the size of the input data. Also, BN and ReLU activation functions are applied to the outputs of the linear layer. GAP blocks are very robust to spatial translations of the inputs as they sum out the spatial information [33]. They also prune the shape of the inputs so that the outputs can be directly fed into the prediction blocks.

The features learned from the feature extraction block are then fed into the scenarios prediction block for the prediction of the five potential scenarios. The input sizes for the three fully connected layers of the block are 256, 128, and 32. Similar to the feature extraction block, both BN and the ReLU activation function are applied for improved convergence of the network. The predicted channel scenario is identified by choosing the output neuron with the largest value.

Given the predicted scenario, the network model then predicts the value of the specific attribute relevant to that scenario, e.g., the distance between DIMMs, transceiver height, and so on. The predicted scenario is one-hot encoded with five bits and then concatenated with the extracted features as the inputs of the attributes prediction block to predict the attribute of that scenario. One-hot encoding translates the integer represented predicted scenario into a group of bits with each bit representing a potential scenario and the predicted class high (1) and others low (0). Since those potential scenarios are independent of each other, one-hot encoding ensures that there is not any presumed hierarchy or partial correspondence between the potential scenario inputs that may be an issue if

of a typical residual block is shown in Fig. 6. As shown in the figure, the input is denoted by x , the mapping of the stacked nonlinear layers in the dotted-line box which consists of a weight layer (convolutional layer), a ReLU activation layer, and another weight layer at the end is denoted as $f(x)$, and a shortcut connection performs the identity mapping to generate the desired output mapping $f(x) + x$. If the input x is already optimal, the weights and biases of the layers $f(x)$ may be optimized to generate a zero output. With the shortcut connection, the network is easier to optimize as it guarantees that the deeper model would not generate higher training error

multiple scenarios were nonzero. Similar to the structure of the scenario prediction block, the attribute prediction block also consists of three fully connected layers whose input sizes are 256, 128, and 32. In contrast with the scenario prediction block, the attribute of the scenario is predicted directly by the block.

Since the proposed network model predicts both the measurement scenario and an attribute of that scenario, training requires both a scenario classification loss and an attribute regression loss. Hence, the overall training loss function of the proposed model is derived as the sum of the cross-entropy loss for the scenarios classification and the mean squared error (MSE) loss for the attribute regression. The cross-entropy loss calculates the expected loss value for the predicted scenario class, while MSE quantifies the distance between the real and predicted value of the attribute. To train the network, we use the robust “Adam” variant of the stochastic gradient descent optimization algorithm which estimates the first- and second-order moments of the gradient via an exponential moving average to update network parameters [38].

IV. ANALYSIS AND MODEL VERIFICATION

In this section, we first trained the proposed ResNet-based feature concatenated network model described in Section III on the augmented training PDP. We then tested and compared the performance of the proposed model with an MLP-based network model on both the expanded measured dataset (without interpolations) and the averaged interpolated dataset (with interpolations) described in Section II-B. Finally, the ResNet- and MLP-based feature concatenated network models are analyzed with Grad-CAM and FAM, respectively, which provide the visual explanations for model decisions.

A. Model Training and Performance Evaluation

Before training, we first standardize the dataset by subtracting the mean and dividing by the standard deviation across the feature dimension. We also normalize the attributes for each scenario. These normalizations allow the network to learn weights on a similar scale for each scenario and attribute. The ResNet-based feature concatenated network model was trained for 1500 epochs. During each epoch, the network iterates over random batches of 256 samples and updates the variables in convolutional and linear layers with each batch. The loss of the model on training, validation, and testing dataset during the training process are compared in Fig. 7. It can be seen from the plot that the Loss of the model on the training dataset is always lower than the loss on the validation and testing dataset. The losses dropped quickly in the first 400 epochs of training as the model parameters are initialized randomly. The losses on the training, validation, and testing datasets converged to 5×10^{-5} , 6.5×10^{-4} , and 5.8×10^{-4} after 1200 epochs of training.

To evaluate the performance of the proposed ResNet-based feature concatenated network model, we first tested it on the expanded measured dataset. The results show that the model can perfectly classify the scenarios under which the input PDP was measured. As described in Section III, our proposed model

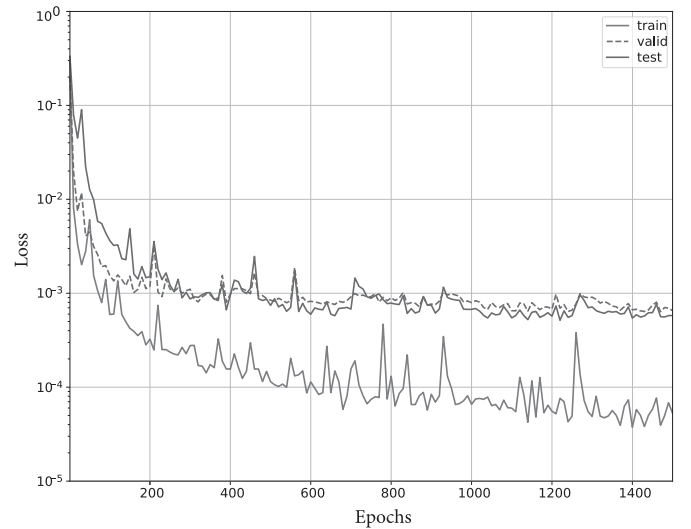


Fig. 7. Losses of ResNet-based feature concatenated network model on training, validation, and testing dataset over 1500 epochs of training.

predicts the attribute under the predicted scenario directly. As with any regression problem, there is always a deviation between the predicted and the correct value. For a robust model, that deviation is small so that the predicted value can be classified as the correct value by choosing an appropriate threshold. To evaluate the robustness of our proposed model, we calculated and compared the classification accuracy of the predicted attribute for each potential scenario with respect to a threshold in the range of zero to half of the step increment size set for the measurements. The maximum distance threshold is 0.15, 0.5, 0.5, 0.125, and 0.5 cm for LoS, RNLoS, OLoS, NLoS, and CPU-PCI link, respectively, as shown in Fig. 8(a). It can be seen from the plot that the ResNet-based feature concatenated model performs perfectly on the expanded measured dataset with achieved nearly 100% classification accuracy under all potential scenarios with the largest possible threshold. Also, it can be observed that the model differentiates between classes most easily in the RNLoS scenario as there are only two classes (either flat or component side of the DIMM as the reflection surface). Since there are only 5 different attributes for the NLoS and CPU-PCI link scenarios, the model differentiates between classes even with relatively small thresholds. Also, comparing these two scenarios, the model converges slower under the CPU-PCI link as the environment of this scenario is much more complicated than that of the NLoS. The small distances between the 16 different attributes for LoS result in slightly less robust performance across a range of thresholds. The model converges the slowest under the OLoS scenario as the measured PDP are very similar to each other with only small variations on the intensity of multipath signals.

The performance of our proposed ResNet-based feature concatenated network model has been compared with an MLP-based method, which is a widely used channel prediction scheme [12], [17], [21], [39], [40], [41], [42], [43], [44], [45], [46], [47]. The structure of the MLP-based network

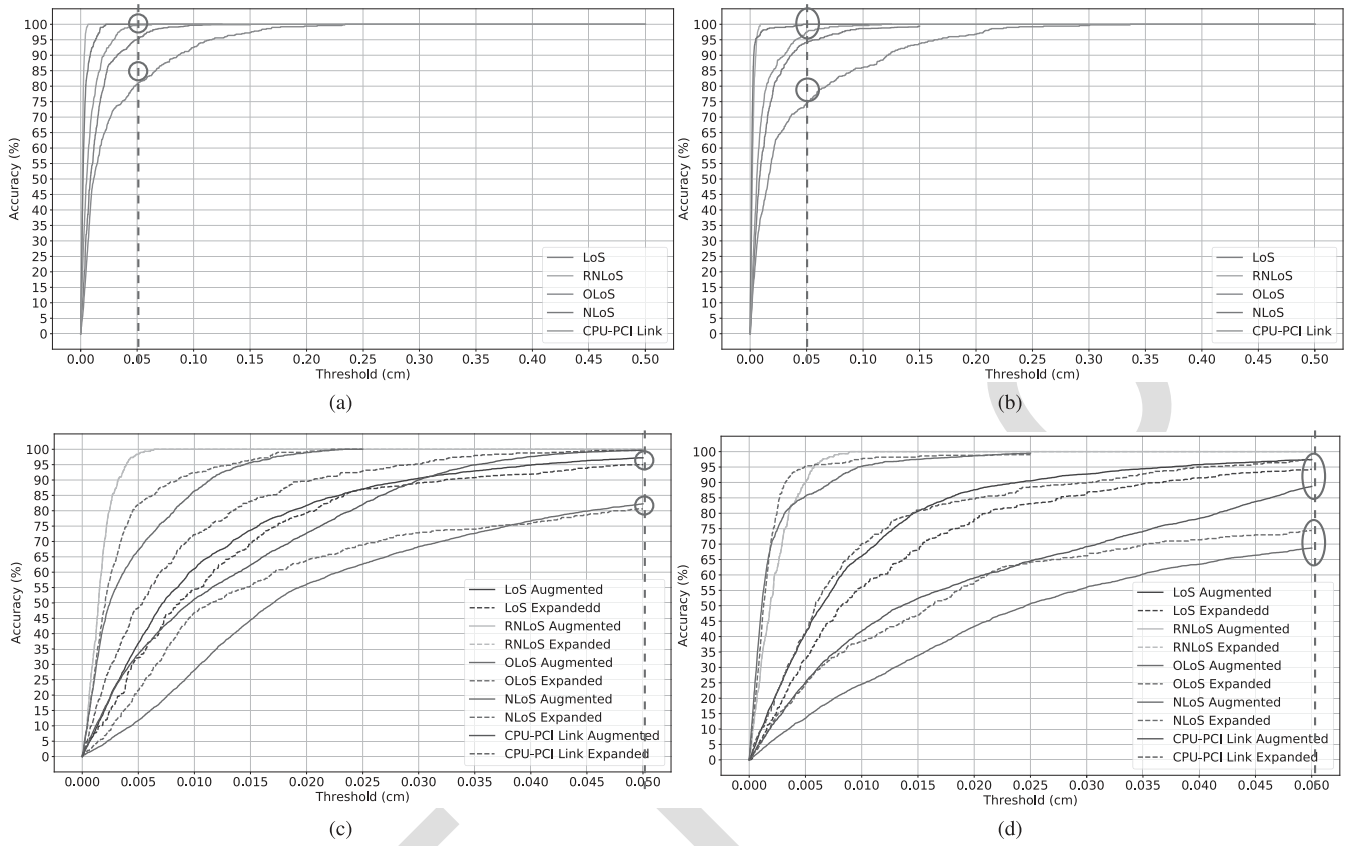


Fig. 8. Classification accuracy of (a) ResNet-based feature concatenated model tested on the expanded measured dataset, (b) MLP-based model tested on the expanded measured dataset, (c) ResNet-based feature concatenated model tested on the averaged interpolated dataset, and (d) MLP-based model tested on the averaged interpolated dataset for the attribute prediction under each potential scenario with respect to the selected threshold.

model used in comparison consists of a five-layer feature extraction block, a three-layer scenario prediction block, and another three-layer attribute prediction block, as shown in Fig. 9. Like our ResNet-based feature concatenated model, the MLP-based model also performs perfectly on the scenario classification task. Fig. 8(b) shows the classification accuracy of the ResNet-based feature concatenated model on the attribute prediction for each potential scenario with respect to the classification threshold on the expanded measured dataset. It can be seen from the plot that the MLP-based model also performs perfectly for the attribute prediction with nearly 100% achieved classification accuracy under all potential scenarios. However, by comparing Fig. 8(a) with (b), it can be observed that our ResNet-based feature concatenated model converges with tighter thresholds than the MLP-based model. Additionally, the classification accuracy achieved by the ResNet-based feature-concatenated model is higher than that of the MLP-based model on LoS, OLoS, and CPU-PCI link with a selected threshold of 0.05 cm, which indicates that the ResNet-based feature-concatenated model is more robust and performs better than the MLP-based model.

In addition to the evaluation on the expanded measured dataset, we also tested and compared our proposed ResNet-based feature concatenated model with the MLP-based model on the averaged interpolated dataset with higher resolution. Again, both models achieve 100% accuracy for scenario classification. However, while the ResNet-based

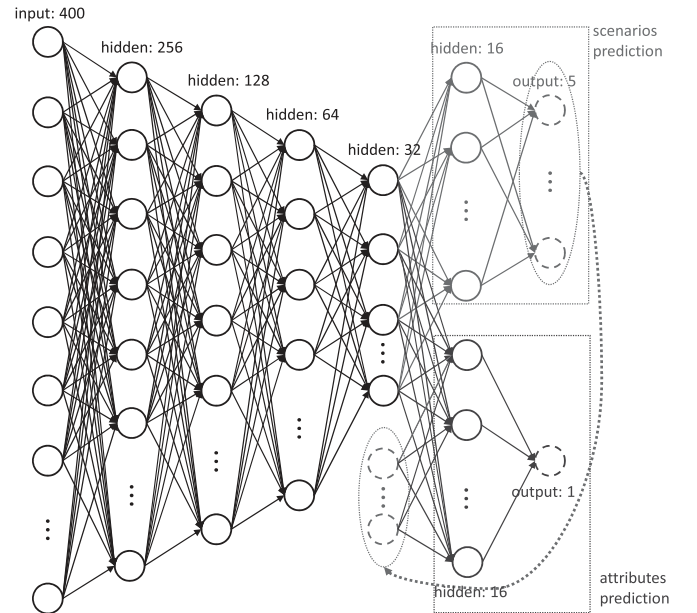


Fig. 9. Structure of the MLP-based network model.

feature concatenated model keeps constant performance for the prediction of the attribute under each scenario, the MLP-based model suffers a performance degradation on the averaged

interpolated dataset. Fig. 8(c) and (d) show the classification accuracy of our ResNet- and the MLP-based feature concatenated network models, respectively, for the attribute prediction under each potential scenarios with respect to the growing threshold on the averaged interpolated dataset and compare with the performances of the models on the measured expanded dataset. The selected threshold is in the range of zero to half of the higher-resolution step increment, which is 0.025 cm for NLoS and 0.05 cm for other scenarios. As shown in Fig. 8(c), it can be seen that the convergence rates of our proposed ResNet-based feature concatenated model on the expanded measured dataset (lower resolution) are very close to those on the averaged interpolated dataset (higher resolution), except for the CPU-PCI link which may because this scenario is much more complicated than others. With the selected threshold of 0.05 cm, as circled in red in these two plots, the classification accuracy of the ResNet base feature concatenated model tested on the averaged interpolated dataset is the same as that of the model tested on the expanded measured dataset. In contrast, it can be noticed from Fig. 8(d) that the convergence rates of the MLP-based model drop a lot on the augmented dataset for nearly all the scenarios, especially for the OLoS and CPU-PCI link. In addition, with the threshold of 0.05 cm, a 9% and a 5% drop of the classification accuracy for OLoS and CPU-PCI link, respectively, can be observed from the plots as circled in red. The performance degradation of the MLP-based model suggests that the range of the deviation between the MLP predicted and the correct value is larger than 0.05 for most attributes under the OLoS and CPU-PCI link, and the MLP-based model performs worse on the averaged interpolated data as the distance between interpolated neighbors are relatively short (half λ for NLoS and λ for others) and these neighbors are very similar to each other. The performance difference of our proposed ResNet- and MLP-based feature concatenated network models indicates that our proposed ResNet-based feature concatenated model is more robust and reliable on the dataset with higher resolution.

B. Model Analysis

To explain the performance difference of these two models and have a better understanding of the wireless channel and the deep learning models, FAM and Grad-CAM are applied on the MLP- and our proposed ResNet-based feature concatenated network models, respectively. To understand how the MLP-based network model makes predictions, we visualize a feature activation map by taking the absolute value of the average over the 256 neurons in the first linear layer as follows:

$$|\mathbf{w}_{\text{ave}}| = \left| \frac{1}{256} \sum_{i=1}^{256} \mathbf{w}_{i,*} \right| \quad (1)$$

where $\mathbf{W} \in R^{256 \times 400}$ represents the weights of the first linear layer [25]. FAM provides an illustration of which parts of the input are emphasized by an average neuron in the first layer to make decisions. As shown in Fig. 10, we highlighted a PDP measured under the OLoS scenario with the calculated $\mathbf{w}_{\text{ave}}$. It can be seen from the plot that the model focuses on the peaks with the excess delay of 0, 1.835, and 3.837 ns to make

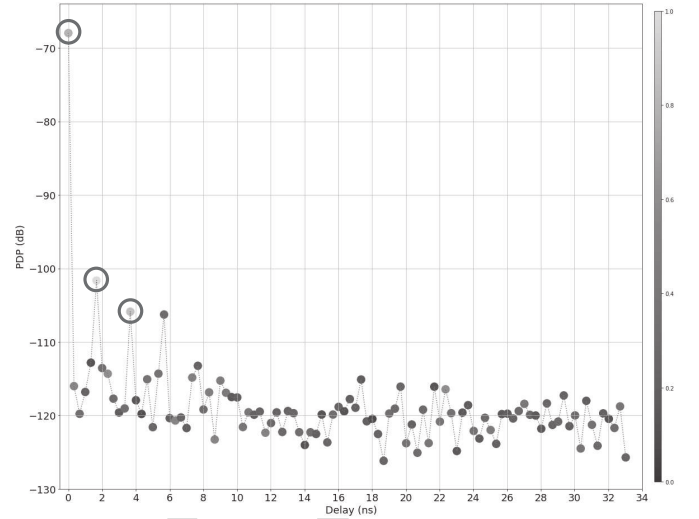


Fig. 10. Contributing regions of the PDP measured under the OLoS scenarios for the MLP-based network model where the attention of the neurons as measured by the FAM is denoted by the color of each point.

decisions, which are circled in red. Instead of the first peak (generated from the signal traveling directly from Tx to Rx) with the most power intensity, the model focuses more on the second and the third peaks (generated from the signal bouncing back and forth between the transceiver sidewalls [16]), which indicates that the model weights more on the multipath signals. This follows intuition since varying the distance between the parallel DIMMs under the OLoS scenario leaves the intensity of the direct traveling signal unchanged while the multipath signal intensity would fluctuate due to interactions of the DIMMs with the Fresnel zones [16]. In spite of this, the intensity of the first peak still influences the model decisions as it has strong differences when we are between different scenarios. Since the intensities of these peaks where the MLP model focuses are very different for the PDP under each scenario [16], the MLP-based model works perfectly on both the expanded measured and averaged interpolated datasets for the scenarios classification task.

The feature map also explains the performance degradation of the MLP-based model on the averaged interpolated dataset. Fig. 11(a) compares the PDP under OLoS scenario with the attributes from 5.2 to 5.6 cm and the part of the confusion matrix of the MLP-based model at these locations. It can be seen from the confusion matrix that the MLP-based model fails in these locations as the classification accuracy is less than 20%. Also, it can be seen that the PDP are very similar and the intensities of the peaks on which the model focuses, as circled in red in the figure, are very close to each other; which explains the failure of the model. As a counterexample, we also plot and compare the PDP with the attributes from 2.0 to 2.3 cm and the part of the confusion matrix at these locations as shown in Fig. 11(b). This figure shows that the MLP-based model performs much better at these locations as the classification accuracy is greater than 80%. It can be observed that the PDP is more easily differentiable between those differences at the peak locations where the model focuses. The contributing

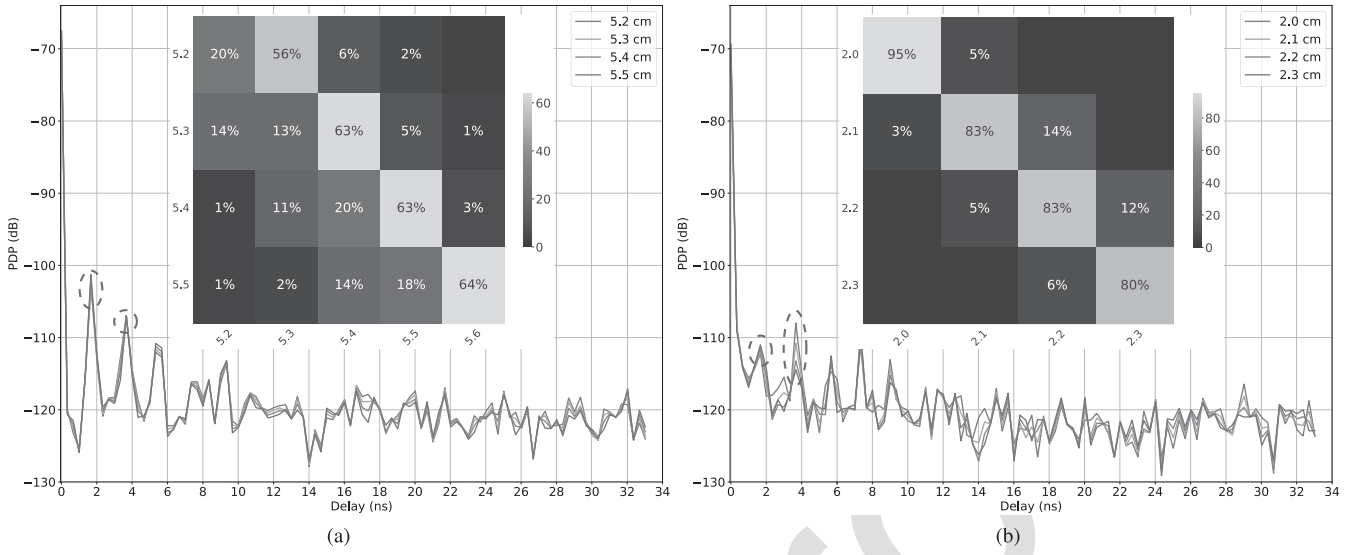


Fig. 11. PDP and confusion matrix of the MLP model under OLoS scenario with (a) attributes from 5.2 to 5.5 cm and (b) attribute from 2 to 2.3 cm.

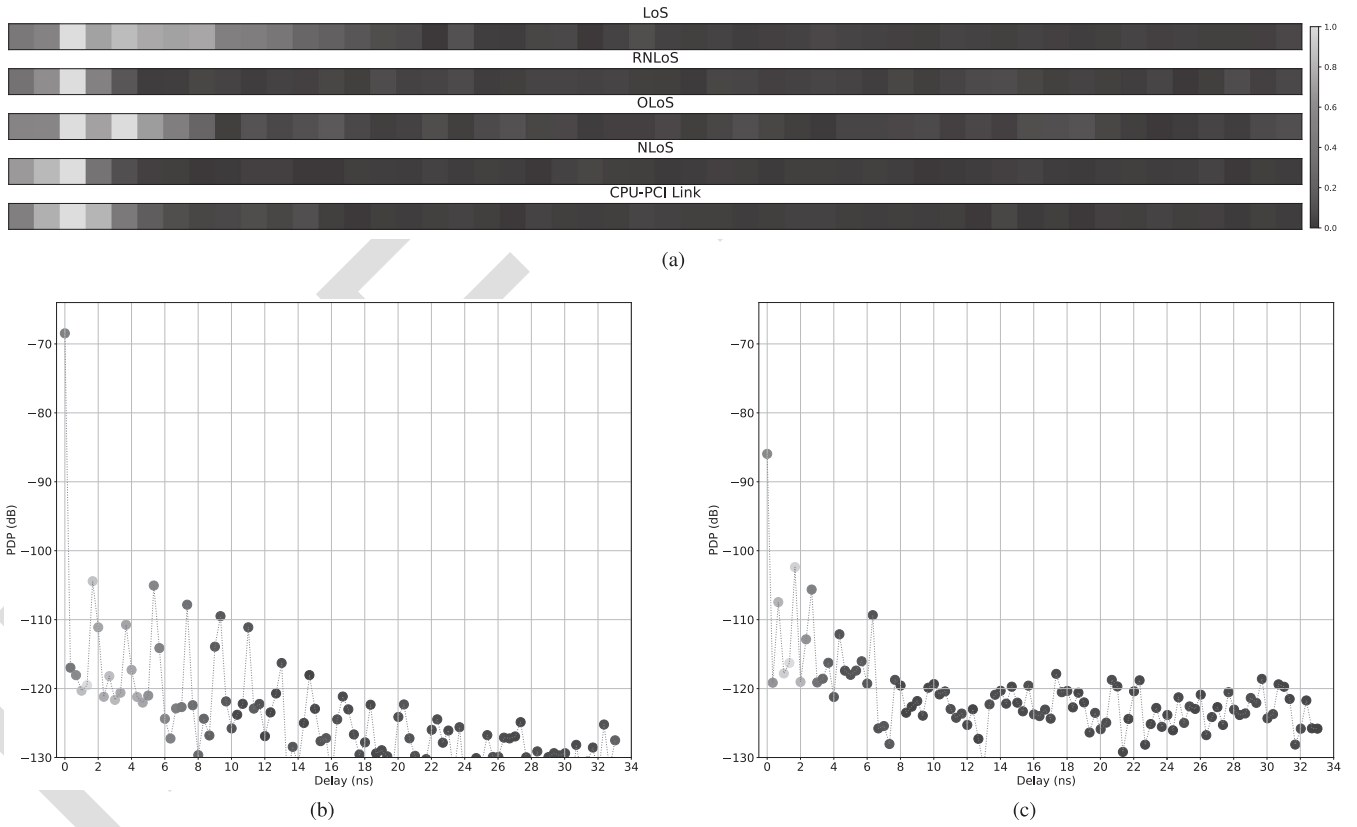


Fig. 12. (a) Heatmap of calculated $L_{Grad-CAM}^c$ with randomly selected PDP as the input for each scenario. (b) Contributing regions colored onto the randomly-selected PDP for the LoS channel. (c) Contributing regions of the randomly-selected PDP for the CPU-PCI link.

regions shown in Fig. 10 also apply to other potential scenarios as the weights of the neurons are fixed for the well trained MLP-based model.

Similarly, we also want to understand how the ResNet-based feature concatenated model makes predictions. However, the FAM is not suitable here as the ResNet-based feature concatenated model has a very different structure than the MLP-based model where convolutional layers instead of fully connected

layers are used to extract the features from the input data. Class activation mapping (CAM) was first proposed to produce a localization map for a classification CNN with a specific architecture, where global average pooled convolutional feature maps are fed directly to the output layer, by projecting back the weights of the output layer on to the convolutional feature maps [48]. A CAM can localize the class-distinguishing region of an input without positional supervision nor the requirement

of a backward pass. However, the constraint on the model architecture is restrictive so that it may not be useful for our case as the GAP layer is connected by another two fully connected networks (scenarios and attributes prediction block) in which the spatial information retained by previous convolutional layers is lost. The Grad-CAM was proposed to address the disadvantage of CAM as it uses the gradient information from the last convolutional layer of the CNN to assign importance values to each neuron for a particular class of interest [26]. The Grad-CAM is used here to provide visual explanations for the decisions made by our ResNet-based feature concatenated model. To derive the class-discriminative localization map $L_{Grad-CAM}^c$ for any class c , the gradient of the score for class c , y^c , with respect to feature map activations A^k of the last convolutional layer are first calculated, $(\partial y^c / \partial A^k)$. These gradients are then global average pooled over the last dimensions (indexed by i) to generate the neuron importance weights α_k^c as follows [26]:

$$\alpha_k^c = \frac{1}{Z} \sum_i \frac{\partial y^c}{\partial A_i^k} \quad (2)$$

where Z is the number of entries in the feature map ($Z = \sum_i 1$). The weight α_k^c captures the importance of feature map k for a target class c . The localization map $L_{Grad-CAM}^c$ is derived by combining weighted forward activation maps and taking the absolute value of it as follows:

$$L_{Grad-CAM}^c = \left| \sum_k \alpha_k^c A^k \right| \quad (3)$$

where K is the number of feature maps of the last convolutional layer. This result provides a coarse heatmap of the same size as the convolutional feature maps in the last convolutional layer. To highlight the regions where the model has more attention to make decisions, we first enlarged the localization map $L_{Grad-CAM}^c$ to the size of input PDP and highlighted the PDP with the enlarged $L_{Grad-CAM}^c$ for each scenario. Instead of using ReLU for the Grad-CAM localizations as described in [26], we take the absolute value as we are interested in the features that have any influence (not just positive influence) on the class of interest. It has been proven that Grad-CAM generalizes CAM for CNN-based architectures as the class features weights ω_k^c used in CAM are proportional to the neuron importance weights α_k^c used by Grad-CAM [26]. This generalization allows us to provide a visual explanation for our ResNet-based feature concatenated model which has complicated structures connected behind the GAP layer. Fig. 12(a) shows the heatmap of the calculated $L_{Grad-CAM}^c$ with randomly selected PDP as the input for each potential scenario. Fig. 12(b) and (c) highlight the contributing regions of the randomly selected PDP for LoS, and CPU-PCI link, respectively, based on the corresponding $L_{Grad-CAM}^c$. It can be observed from the plots that, similar to the MLP-based model, the ResNet-based feature concatenated model also focuses on the first few peaks of the PDP with specific emphasis depending on the scenario. Interestingly, both models pay more attention to the multipath signals than the direct traveling signal. However, besides the peaks,

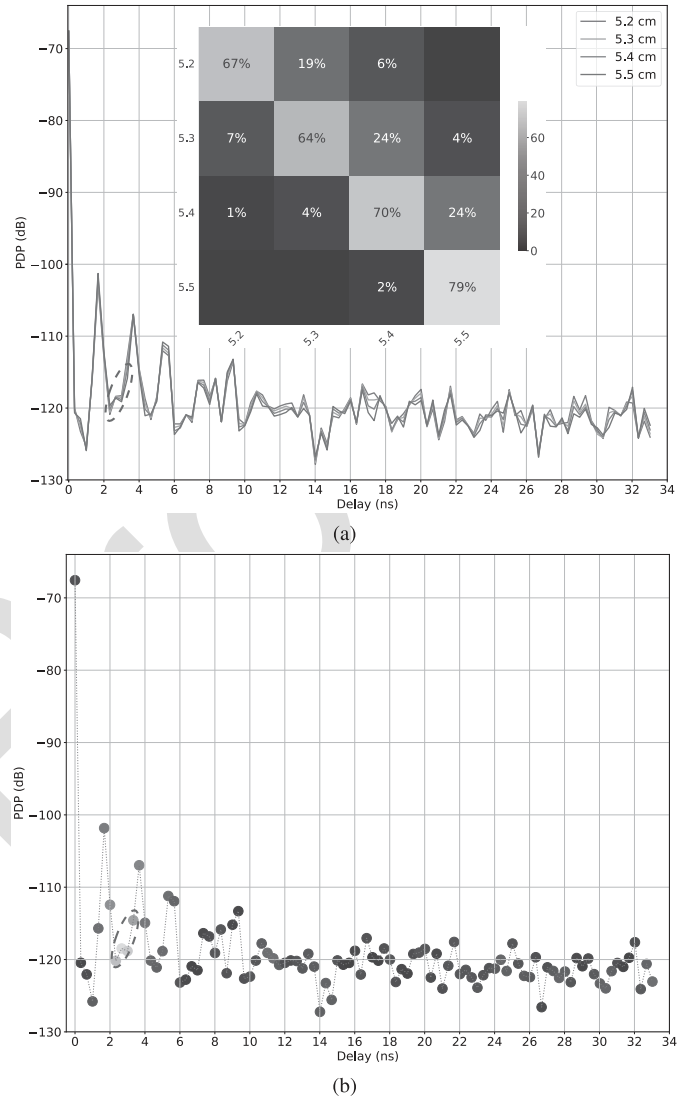


Fig. 13. (a) PDP and confusion matrix of the ResNet-based feature concatenated model under OLoS scenario with attributes from 5.2 to 5.5 cm. (b) Highlighted contributing regions of the PDP with the attribute of 5.3 cm under OLoS scenario.

the ResNet-based feature concatenated model also focuses on the valleys between those peaks even with stronger attention. This indicates that these valleys can also provide important information on the wireless channel and could also explain why the ResNet-based feature concatenated model performs more robustly than the MLP model. Fig. 13(a) shows the PDP and the confusion matrix of the ResNet-based feature concatenated model, again, under the OLoS scenario with the attribute ranging from 5.2 to 5.5 cm. Compared with the confusion matrix of the MLP model shown in Fig. 11(a), it can be concluded that the ResNet-based feature concatenated model performs much better than the MLP model at these locations as the classification accuracy of the ResNet-based feature concatenated model is greater than 64% while that of the MLP model is even less than 20%. Fig. 13(b) shows the contributing regions of the PDP with the attribute of 5.3 cm under the OLoS scenario. As shown in

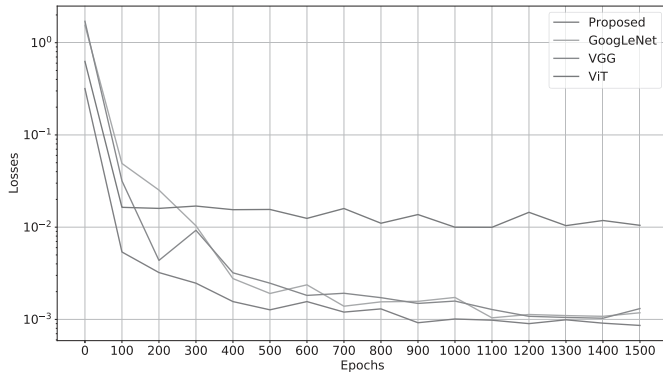


Fig. 14. Testing losses of our proposed ResNet-based feature concatenated model, GoogLeNet, VGG, and ViT on the same testing dataset throughout the training process.

this plot, besides the peaks due to the multipath signals, the model has even more attention on the valleys between those peaks where the PDP with the attributes from 5.2 to 5.5 cm can be clearly distinguished with each other as circled in red in Fig. 13(a), which explains why the ResNet-based feature concatenated model performs better at these locations.

Our proposed model performs better than other models as it has strong attention on the valleys between the multipath peaks. To show this, we compared the performance of our proposed ResNet-based feature concatenated model with some other mainstream deep learning models which include the GoogLeNet, visual geometry group (VGG), and vision transformer (ViT). Fig. 14 shows the testing losses of our proposed model, GoogLeNet, VGG, and ViT when evaluated on the same testing dataset throughout the training process. As shown in the dataset, our ResNet-based feature concatenated model outperforms other models, which illustrates the superiority of our proposed model.

V. CONCLUSION

This article presents the application of deep learning method on property prediction of the THz wireless channel in a motherboard desktop environment. Measurements were performed on a motherboard in a desktop size metal cavity. The PDP data which contains the channel information fed into the deep learning models was collected under five environment scenarios. For each scenario, measurements were taken several times with different attributes. To train the proposed deep learning model robustly and evaluate their performance comprehensively, data augmentation is applied on the measured PDP to expand the data size and to improve the resolution. A ResNet-based feature concatenated neural network model was proposed to perform the channel prediction. The generalized model structure consists of three blocks which are feature extraction, scenarios prediction, and attributes prediction block. The proposed model is evaluated and compared with an MLP-based model on both expanded measured and averaged interpolated datasets. Both models perform well on the expanded measured dataset with nearly perfect prediction

on both scenarios and attributes. Our proposed model outperforms the MLP-based model on the averaged interpolated dataset as the MLP model has a drop of classification accuracy on the attribute prediction task for each scenario with a fixed threshold and converges slower with respect to the larger thresholds, while our ResNet-based feature concatenated model keeps constant performance. The analyses are made with FAM and Grad-CAM applied on the MLP- and our ResNet-based feature concatenated models, respectively. It is shown from the FAM that the MLP model has a strong focus on the peaks of the PDP corresponding to multipath signals. However, some of the adjacent interpolated data are very close to each other in the regions on which the MLP focuses, which leads to performance degradation (or failure) on these areas. Also, as seen from the Grad-CAM, the ResNet-based feature concatenated model also has strong attention on the valleys between peaks where the neighboring interpolated data can be distinguished better, which accounts for its superior performance. Better performance in comparison with other mainstream deep learning models also proves the superiority of our proposed ResNet-based feature concatenated model.

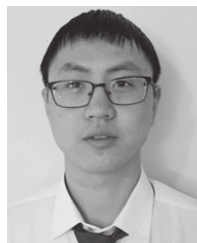
ACKNOWLEDGMENT

The views and findings in this paper are those of the authors and do not necessarily reflect the views of NSF.

REFERENCES

- [1] J. Fu, P. Juyal, and A. Zajic, "THz channel characterization of chip-to-chip communication in desktop size metal enclosure," *IEEE Trans. Antennas Propag.*, vol. 67, no. 12, pp. 7550–7560, Dec. 2019.
- [2] S. Kim and A. Zajic, "Characterization of 300-GHz wireless channel on a computer motherboard," *IEEE Trans. Antennas Propag.*, vol. 64, no. 12, pp. 5411–5423, Dec. 2016.
- [3] S. Kim and A. Zajic, "Statistical modeling and simulation of short-range device-to-device communication channels at sub-THz frequencies," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6423–6433, Sep. 2016.
- [4] J. Fu, P. Juyal, and A. Zajic, "Modeling of 300 GHz chip-to-chip wireless channels in metal enclosures," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3214–3227, May 2020.
- [5] J. Fu, P. Juyal, B. B. Yilmaz, and A. Zajic, "Investigation of resonance based propagation loss modeling for THz chip-to-chip wireless communications," in *Proc. 14th Eur. Conf. Antennas Propag. (EuCAP)*, 2020, pp. 1–5.
- [6] C.-L. Cheng, S. Sangodoyin, and A. Zajic, "THz cluster-based modeling and propagation characterization in a data center environment," *IEEE Access*, vol. 8, pp. 56544–56558, 2020.
- [7] G. P. Fettweis, N. Ul Hassan, L. Landau, and E. Fischer, "Wireless interconnect for board and chip level," in *Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE)*, 2013, pp. 958–963.
- [8] T. Kurner, A. Fricke, S. Rey, P. Le Bars, A. Mounir, and T. Kleine-Ostmann, "Measurements and modeling of basic propagation characteristics for intra-device communications at 60 GHz and 300 GHz," *J. Infr., Millim., Terahertz Waves*, vol. 36, no. 2, pp. 144–158, Feb. 2015.
- [9] A. Fricke, M. Achir, P. Le Bars, and T. Kurner, "A model for the reflection of terahertz signals from printed circuit board surfaces," *Int. J. Microw. Wireless Technol.*, vol. 10, no. 2, pp. 179–186, Mar. 2018.
- [10] J. Fu, "THz wireless channel characterization and modeling for chip-to-chip communication in computing systems," Ph.D. dissertation, Georgia Inst. Technol., Atlanta, GA, USA, 2022.
- [11] IT Association. *Infiniband Roadmap*. [Online]. Available: <https://www.infinibandta.org/infiniband-roadmap/>

- [12] C. Huang et al., "Artificial intelligence enabled radio propagation for communications—Part II: Scenario identification and channel modeling," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3955–3969, Jun. 2022.
- [13] D. He et al., "Stochastic channel modeling for kiosk applications in the terahertz band," *IEEE Trans. Terahertz Sci. Technol.*, vol. 7, no. 5, pp. 502–513, Sep. 2017.
- [14] R. Piesiewicz, C. Jansen, D. Mittleman, T. Kleine-Ostmann, M. Koch, and T. Kurner, "Scattering analysis for the modeling of THz communication systems," *IEEE Trans. Antennas Propag.*, vol. 55, no. 11, pp. 3002–3009, Nov. 2007.
- [15] S. Priebe, M. Jacob, C. Jansen, and T. Kurner, "Non-specular scattering modeling for THz propagation simulations," in *Proc. 5th Eur. Conf. Antennas Propag. (EUCAP)*, 2011, pp. 1–5.
- [16] J. Fu, P. Juyal, and A. Zajic, "Generalized modeling and propagation characterization of THz wireless links in computer desktop environment," *Radio Sci.*, vol. 57, no. 4, pp. 1–26, Apr. 2022.
- [17] J. Fu, P. Juyal, E. J. Jorgensen, and A. Zajic, "Comparison of statistical and deep learning path loss model for motherboard desktop environment," in *Proc. 16th Eur. Conf. Antennas Propag. (EuCAP)*, Mar. 2022, pp. 1–5.
- [18] C. Huang et al., "Artificial intelligence enabled radio propagation for communications—Part I: Channel characterization and antenna-channel optimization," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3939–3954, Jun. 2022.
- [19] M. Uccellari et al., "On the application of support vector machines to the prediction of propagation losses at 169 MHz for smart metering applications," *IET Microw., Antennas Propag.*, vol. 12, no. 3, pp. 302–312, Feb. 2018.
- [20] J. Salmi, A. Richter, and V. Koivunen, "Detection and tracking of MIMO propagation path parameters using state-space approach," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1538–1550, Apr. 2009.
- [21] Y.-T. Ma, K.-H. Liu, and Y.-N. Guo, "Artificial neural network modeling approach to power-line communication multi-path channel," in *Proc. Int. Conf. Neural Netw. Signal Process.*, Jun. 2008, pp. 229–232.
- [22] Q. Zheng et al., "Channel non-line-of-sight identification based on convolutional neural networks," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1500–1504, Sep. 2020.
- [23] M. I. AlHajri, N. T. Ali, and R. M. Shubair, "Classification of indoor environments for IoT applications: A machine learning approach," *IEEE Antennas Wireless Propag. Lett.*, vol. 17, no. 12, pp. 2164–2168, Dec. 2018.
- [24] M. Yang et al., "Machine-learning-based scenario identification using channel characteristics in intelligent vehicular communications," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 3961–3974, Jul. 2021.
- [25] E. J. Jorgensen, F. T. Werner, M. Prvulovic, and A. Zajic, "Deep learning classification of motherboard components by leveraging EM side-channel signals," *J. Hardw. Syst. Secur.*, vol. 5, no. 2, pp. 114–126, Jun. 2021.
- [26] R. R. Selyaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.
- [27] Virginia Diodes. *Custom Transmitters*. [Online]. Available: <https://vadiodes.com/en/products/custom-transmitters>
- [28] UI Communications. *Phase Locked Dielectric Resonator Oscillator*. [Online]. Available: <https://www.ultra.group/media/2081/ultra-herley-series-pdro.pdf>
- [29] Millimeter. *Frequency Multiplier—33 to 50 GHz*. [Online]. Available: <https://nordengroup.com/product-group/frequency-multiplier-33-to-50-ghz/>
- [30] Virginia Diodes. *Multipliers—VDI Model: WR6.5X3*. [Online]. Available: <https://www.vadiodes.com/en/wr6-5x3>
- [31] Virginia Diodes. *Mixers—VDI Model: WR2.8SHM*. [Online]. Available: <https://www.vadiodes.com/en/wr2-8shsm>
- [32] S. Kim and A. G. Zajic, "A path loss model for 300-GHz wireless channels," in *Proc. IEEE Antennas Propag. Soc. Int. Symp. (APSURSI)*, Jul. 2014, pp. 1175–1176.
- [33] M. Lin, Q. Chen, and S. Yan, "Network in network," 2013, *arXiv:1312.4400*.
- [34] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*.
- [35] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. 14th Int. Conf. Artif. Intell. Statist.*, 2011, pp. 315–323.
- [36] M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex," *Nature Neurosci.*, vol. 2, no. 11, pp. 1019–1025, Nov. 1999.
- [37] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [39] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, Jan. 1989.
- [40] M. Kalakh, N. Kandil, and N. Hakem, "Neural networks model of an UWB channel path loss in a mine environment," in *Proc. IEEE 75th Veh. Technol. Conf. (VTC Spring)*, May 2012, pp. 1–5.
- [41] N. Zaarour, N. Kandil, N. Hakem, and C. Despins, "Comparative experimental study on modeling the path loss of an UWB channel in a mine environment using MLP and RBF neural networks," in *Proc. Int. Conf. Wireless Commun. Underground Confined Areas*, Aug. 2012, pp. 1–6.
- [42] N. Zaarour, N. Kandil, and N. Hakem, "An accurate neural network approach in modeling an UWB channel in an underground mine," in *Proc. IEEE Antennas Propag. Soc. Int. Symp. (APSURSI)*, Jul. 2013, pp. 1608–1609.
- [43] N. Zaarour, S. Affes, N. Kandil, and N. Hakem, "Comparative study on a 60 GHz path loss channel modeling in a mine environment using neural networks," in *Proc. IEEE Int. Conf. Ubiquitous Wireless Broadband (ICUWB)*, Oct. 2015, pp. 1–4.
- [44] W. Hu, S. Geng, and X. Zhao, "Mm-wave 60 GHz channel fading effects analysis based on RBF neural network," in *Proc. 5th Int. Conf. Comput. Commun. Syst. (ICCCS)*, May 2020, pp. 589–593.
- [45] A. Neskovic, N. Neskovic, and D. Paunovic, "Macrocell electric field strength prediction model based upon artificial neural networks," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1170–1177, Aug. 2002.
- [46] J. A. Romo, I. F. Anitzine, and F. P. Fontan, "Application of neural networks to field strength prediction for indoor environments," in *Proc. 1st Eur. Conf. Antennas Propag.*, Nov. 2006, pp. 1–6.
- [47] H. Zhou, F. Wang, and C. Yang, "Application of artificial neural networks to the prediction of field strength in indoor environment for wireless LAN," in *Proc. Int. Conf. Wireless Commun., Netw. Mobile Comput.*, 2005, pp. 1189–1192.
- [48] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2921–2929.



Jinbang Fu was born in Nanchang, China. He received the B.Sc. degree (summa cum laude) in electrical and computer engineering from the University of Massachusetts, Amherst, MA, USA, in 2017. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the Georgia Institute of Technology, Atlanta, GA, USA.

He is a Graduate Research Assistant with the Electromagnetic Measurements in Communications and Computing Laboratory, Georgia Institute of Technology. His current research interests include THz wireless channel measurements and modeling and THz wireless communication system design for chip-to-chip communication.



Erik J. Jorgensen (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from the University of Wisconsin, Madison, WI, USA, in 2017, and the M.Sc. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2019, where he is currently pursuing the Ph.D. degree in machine learning with the Electromagnetic Measurements in Communications and Computing (EMC2) Laboratory.

His current research interests include explainable machine learning, deep learning methods, computer vision, and optimization.



Prateek Juyal received the B.E. degree in electronics and communication engineering from Kumaun University, Nainital, India, in 2007, the M.Tech. degree in digital communication from Guru Gobind Singh Indraprastha University, New Delhi, India, in 2009, and the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Winnipeg, MB, Canada, in 2017.

Currently, he is working as a Post-Doctoral Fellow with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. His current research interests include sensors, security, and wireless communication.



Alenka Zajić (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees from the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia, in 2001 and 2003, respectively, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2008.

Currently, she is an Associate Professor at the School of Electrical and Computer Engineering, Georgia Institute of Technology. Prior to that, she was a Visiting Faculty Member at the School of Computer Science, Georgia Institute of Technology, a Post-Doctoral Fellow at the Naval Research Laboratory, Washington, DC, USA, and a Design Engineer at Skyworks Solutions Inc, Irvine, CA, USA. Her research interests span areas of electromagnetic, wireless communications, signal processing, and computer engineering.

Dr. Zajić was a recipient of the 2017 NSF CAREER award, the Best Paper Award at MICRO 2016, 2012 Neal Shepherd Memorial Best Propagation Paper Award, the Best Student Paper Award at the IEEE International Conference on Communications and Electronics 2014, the Best Paper Award at the International Conference on Telecommunications 2008, the Best Student Paper Award at the 2007 Wireless Communications and Networking Conference, and the Dan Noble Fellowship in 2004, which was awarded by Motorola Inc. and the IEEE Vehicular Technology Society for quality impact on the area of vehicular technology.