ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-7605-5/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICASSP39728.2021.9414388

DEEP CONVOLUTIONAL GAUSSIAN PROCESSES FOR MMWAVE OUTDOOR LOCALIZATION

Xuyu Wang[⋆] *Mohini Patil*[⋆] *Chao Yang*[†] *Shiwen Mao*[†] *Palak Anilkumar Patel*[⋆]

[⋆]Department of Computer Science, California State University, Sacramento, CA 95819, USA
[†]Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849-5201, USA

ABSTRACT

Millimeter Wave (mmWave) communications, as a core technique of 5G, can be leveraged for outdoor localization because of its large bandwidth and massive antenna array. Fingerprinting based mmWave outdoor localization methods using deep learning are highly suitable for non-line-of-sight (NLOS) environments. In this paper, we propose a deep convolutional Gaussian process (DCGP) based regression approach to achieve high robustness for fingerprinting-based mmWave outdoor localization, which exploits the convolutional structure for deep Gaussian process to allow uncertainty estimation on location predictions. Specially, we present a system architecture of mmWave based outdoor localization, including beamforming image construction and DCGP training, where DCGP model can effectively learn the location features from mmWave beamforming images. Our experimental results show that the proposed DCGP method can achieve higher outdoor localization accuracy than a CNN-based baseline method.

***Index Terms*—** Outdoor Localization, 5G mmWave, Beamforming, Convolutional Neural Network, Deep Convolution Gaussian Process.

1. INTRODUCTION

With the rapid growth of wireless communications and networks, new wireless applications such as location based services (LBS) are quickly growing, which can greatly improve user's experience and impact people's lives [1]. Currently, outdoor location information is mostly obtained by the global positioning system (GPS), which can achieve a localization accuracy at about 5 m in line-of-sight (LOS) conditions for civilian use [1]. However, the GPS performance could be poor in rich-scattering environments and urban canyons, and does not work indoors. Alternative outdoor localization techniques such as long term evolution (LTE) and LoRa with received signal strength (RSS) [2] usually does not achieve high outdoor localization accuracy. For example, the median error of network-based localization with LTE is 80 m in urban environments [3].

Millimeter Wave (mmWave) communications, as a core technique of 5G, can not only provide high data rates, but also have high temporal resolution and high directivity. Various mmWave-based localization methods have been proposed, which mainly leverage geometric methods (e.g., angle-of-arrival (AOA) and time-of-arrival (TOA)) for theoretical study of mmWave localization [4]. For example, in LOS environments, fundamental limits of position and rotation angle estimation from a single transmitter in 2D and 3D space are derived in [5]. Moreover, in multipath scenarios, the strongest RSS is regarded as the LOS. However, this assumption will not hold true in outdoor cluttered environments. Recent mmWave experimental work shows that geometry-based methods cannot be directly applied to accurately locate devices in non-line-of-sight (NLOS) environments, because of the non-linear mmWave propagation [6].

Fingerprinting solutions with deep learning have been proposed recently, which is highly suitable for NLOS environments and can achieve better performance than traditional machine learning based schemes. This is because the received signal from NLOS can be used as features for location estimation with deep learning. Our indoor localization works focus on using different deep learning (e.g., deep autoencoder, convolutional neural network (CNN), deep residual sharing learning) with WiFi channel state information (CSI) data [7–9]. There are several recent work on applying deep learning to mmWave fingerprinting. For example, multiple access points are used to create a location fingerprint database of received powers and AOA in [10], and CNN has been applied to mmWave outdoor fingerprinting in [11, 12]. Meanwhile, other machine learning methods have been employed for single user [13, 14] and multi-user [15] indoor localization with commodity 60 GHz mmWave devices.

Motivated by the above works, we study the problem of mmWave fingerprinting based outdoor localization in this paper. Although the existing deep learning based approaches (e.g., CNN) can achieve acceptable outdoor localization accuracy for mmWave networks in some cases, there is no guarantee on their robustness, because of the *black-box approach* of the deep learning models. In this paper, we propose a

This work is supported in part by the NSF under Grants ECCS-1923163 and CNS-1822055 , and the Wireless Engineering Research and Education Center at Auburn University, Auburn, AL

deep convolutional Gaussian process (DCGP) based *regression* approach for mmWave outdoor localization, which is different from the original DCGP method [16] that is focused on image *classification* problems. In addition, compared with CNN, DCGP is a fully Bayesian kernel method with no neural network component, which can provide uncertainty estimation on location predictions [16]. Also, DCGP leverages the convolutional structure for deep Gaussian process, which can detect hierarchical combinations of local features in the mmWave dataset, especially in NLOS environments.

In particular, we present the system design for mmWave-based localization, including beamforming image construction and DCGP training, where the DCGP model can effectively extract the features of mmWave beamforming images. Then, we introduce the mmWave system model and show how to construct mmWave beamforming images. We then examine how to effective train the DCGP model with beamforming images. Finally, we evaluate the localization performance of the proposed scheme, and compare its performance with a CNN-based baseline scheme, using an open-source mmWave dataset for outdoor experiments. The experimental results show that the proposed DCGP method can achieve better localization performance than the CNN-based method.

The remainder of this paper is organized as follows. We discuss the system design in Section 2 and our performance study in Section 3. Section 4 summaries this paper.

2. SYSTEM DESIGN

2.1. System Architecture

The system architecture of mmWave based outdoor localization is shown in Fig. 1. We consider a mobile device that can receive mmWave power delay profile (PDP) data from a base station (BS). The BS can send short pulses with a sequence of beamforming patterns that cover all transmission angles in a measured area. The system needs to create a beamforming image in each location that captures the features of PDP data in mmWave transmission. Then, the collected beamforming images from all locations can be leveraged to train a DCGP in the offline stage, which estimates the mobile device location with a measured beamforming image in the online stage.

In addition, a mmWave ray-tracing simulation on accurate 3D maps can generate the propagation dataset and location labels. Specially. the Wireless InSite ray-tracking software with a highly accurate open-source 3D map in the NYU area is used to generate the outdoor mmWave dataset. Compared with the model-based mmWave localization methods, the proposed method can be more effective in outdoor NLOS cases. In addition, the proposed DCGP can improve the robustness of outdoor localization with beamforming images compared with CNN. Compared with our previous work DeepMap, that uses a deep Gaussian process for indoor location estimation with WiFi RSS data sequence (i.e., 1-D dataset) [17], this work considers 2D beamforming image dataset with DCGP, which is more effective for mmWave outdoor localization.

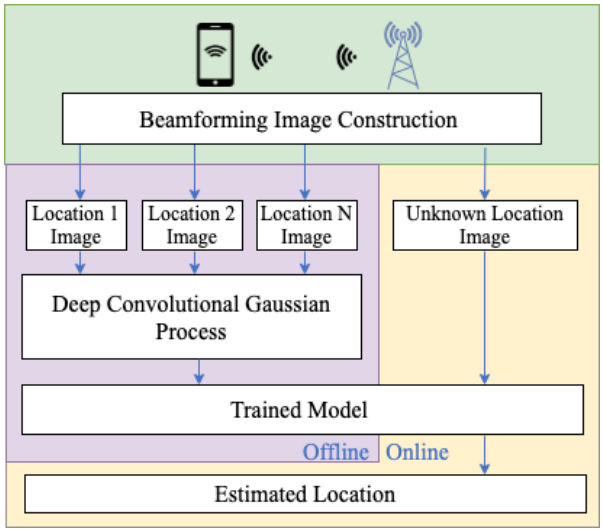

Fig. 1. System architecture of mmWave based localization.

2.2. Constructing Beamforming Images

We focus on data-driven beamforming based mmWave fingerprinting for outdoor localization. Motivated by our prior work on CSI image-based localization [11,17], we also create mmWave beamforming images as fingerprints, which can be effectively handled by the DCGP model [16] to estimate location. DCGP integrates the convolutional structure and Gaussian process in a Bayesian framework, which can effectively extract hierarchical features from 2-D beamforming images to improve the robustness of localization performance. Before we discuss DCGP based localization, we first introduce the mmWave system model and then construct mmWave beamforming images.

We consider a mmWave system with phase-shifter based analog beamforming, which has a low hardware requirement [18]. Assume a fixed codebook C_T with B_T beam patterns for a BS with N_B antennas, which covers all possible angles of mmWave transmissions. The BS can transmit a 28GHz mmWave signal with B_T beam patterns sequentially. We also consider a mobile device with N_R antennas. The received signal y_i at the mobile device for the ith transmitting beam pattern in the frequency-domain is given by

$$y_i = W^T H F_i s + W^T z, \tag{1}$$

where the superscript T represents a matrix transpose, W is the beamforming vector at the mobile device, H is the channel matrix, F_i is the ith transmitting beam pattern, s is the transmitted signal, and z is the additive Gaussian white noise. Because analog beamforming is based on a codebook design, the current beam pattern is $F_i \in C_T = \{F_1, F_2, ..., F_{B_T}\}$.

We assume that the receiver also employs a beamforming codebook C_R with B_R beam patterns to include all AoA values with a similar gain, where $C_R = \{W_1, W_2, ..., W_{B_R}\}$. Because there are B_R beam patterns at the mobile device, it will sample the original transmission B_R times. To create the beamforming image data, only the maximum measured value is stored for each transmitting beam pattern. Thus, the stored data $m_i(n)$ in the nth sample from the ith transmitting beam pattern is calculated as

$$m_i(n) = \max_{k \in \{1,2,...,B_R\}} y_i^k(nT_0), n = 1, ..., N, \quad (2)$$

where $y_i^k(\cdot)$ is the received signal using the ith transmitting beam pattern and the kth receiving beam pattern, and N is the number of samples and T_0 is the sampling period in seconds. Note that for each transmitting beam pattern, the received PDP values can be obtained based on $m_i(n), n = 1, 2, ..., N$. Thus, using all the transmitting B_T beam patterns, we can create a 2-D beamforming image M with the received PDP values in all transmitting beam patterns, which is defined as

$$M = \begin{bmatrix} m_1(1) & m_2(1) & \dots & m_{B_T}(1) \\ m_1(2) & m_2(2) & \dots & m_{B_T}(2) \\ \vdots & \vdots & \ddots & \vdots \\ m_1(N) & m_2(N) & \dots & m_{B_T}(N) \end{bmatrix}. \quad (3)$$

The created 2-D beamforming image is illustrated in Fig. 2, where the x-axis represents the index of the transmitting beam pattern of the BS, and the y-axis represents the sample index of the received PDP in a given beam pattern. Using the Wireless InSite ray-tracking software with the highly accurate 3D map in the NYU area, a dataset is created with 160,801 beamforming images, which will be used to train the DCGP model in the following.

2.3. Training Deep Convolutional Gaussian Process

We use the 2D beamforming images to train the DCGP for location estimation, which iteratively convolutes several Gaussian process (GP) functions over the images. We formulate the mmWave localization problem as a supervised DCGP regression problem with input beamformng image M and a output vector y that represents the 2-D locations, as shown in Fig. 2. The details on DCGP training are discussed below.

Consider an image representation $f_c^l = \mathbb{R}^{H_l \times W_l}$ of width W_l and height H_l (i.e., number of pixels) at layer l for channel c. We then create a 3D tensor $f^l = (f_1^l, f_2^l, ..., f_C^l) \in \mathbb{R}^{H_l \times W_l \times C_l}$ at the lth layer, where C is the number of channels. We have $C_0 = 1$ for the input image at layer 0 using a 2-D beamforming image M generated as before. We could create a sequence of layers f^l to map the input image M to the output y that is $M \to f^1 ... \to f^L = y$, where $L = 3$ for the proposed method. To implement the convolutional operation on the GP layers, a 3D tensor f^l can be divided into

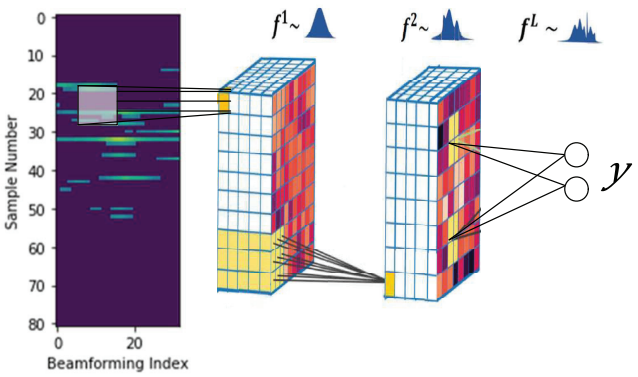

Fig. 2. Training the DCGP model.

several patches $f^l[p] \in \mathbb{R}^{h_l \times w_l \times C_l}$, where p is the patch index, h_l and w_l are the height and width of the patch at the lth layer, respetively.

In addition, we construct the layers by using convolutions of patch response functions $g_c^l : \mathbb{R}^{h_{l-1} \times w_{l-1} \times C_{l-1}} \to \mathbb{R}$ over the input one patch at a time, which can build the next layer representation. The C patch responses at each of the first $L-1$ layers are considered as independent GPs with shared prior [16], where each patch response is defined by

$$g_c^l(f^{l-1}[p]) \sim GP(0, k(f^{l-1}[p], f'^{l-1}[p'])), \; c = 1, 2, ..., C, \quad (4)$$

The kernel $k(\cdot, \cdot)$ represents the similarity of two image patches. We adopt the radial basis function (RBF) kernel in DCGP. For two patches x_1 and x_2, it is defined as [17]

$$k(x_1, x_2) = \sigma_f^2 \exp\left(-\frac{1}{2\gamma^2} |x_1 - x_2|^2\right), \quad (5)$$

where σ_f and γ are the hyper-parameters; σ_f^2 represents the variance and γ is a length scale.

To augment each patch response function, we use a set of I inducing patches z^l in the patch space $\mathbb{R}^{h_{l-1} \times w_{l-1} \times C_{l-1}}$ with the corresponding response r_c^l [16]. Then, we obtain I inducing patches $Z^l = (z_1^l, z_2^l, ..., z_I^l)$ that are shared among the C patch response functions. In fact, each patch response function has different inducing responses $R_c^l = (r_{c1}^l, r_{c2}^l, .., r_{cI}^l)$. The conditional patch response probability is defined by

$$\Pr(g_c^l | f^{l-1}, R_c^l, Z^l) = N(g_c^l; \mu, \Sigma) \quad (6)$$

$$\mu = K_{f^{l-1}Z^l} K_{Z^l, Z^l}^{-1} R_c^l \quad (7)$$

$$\Sigma = K_{f^{l-1}f^{l-1}} - K_{f^{l-1}Z^l} K_{Z^l, Z^l}^{-1} K_{Z^l f^{l-1}}, \quad (8)$$

where $K_{f^{l-1}Z^l}$ is the covariance matrix between the input and the inducing variables that can estimate the similarity of all the patches.

In the last regression layer, a GP with a weighted convolutional kernel is used to aggregate the convolutional layer outputs. We adopt the doubly stochastic variational inference [19] to learn the DCGP model, where all the parameters (i.e., RBF kernel parameters and the patch weights in the last layer) in the DCGP are trained by the Adam optimizer. In the online stage, the trained DCGP model will be used for location predication with a measured beamforming image.

3. EXPERIMENTAL STUDY

3.1. Dataset and System Configuration

In this paper, we use the open-source mmWave dataset using the ray-tracking software (i.e., Wireless InSite 3.0.0.1) at New York university of a 400 $\times$ 400 m^2 area [18]. The dataset includes beamforming images from 160,801 bidimensional positions. The ray-tracking simulation parameters are based on experimental measurements [20]. Specially, the mmWave carrier frequency is set to 28 GHz, the codebook size is 32, the number of samples for each beam pattern at the transmitter is 82, the received PDP is sampled at 20 MHz, the antenna gains of transmitter and receiver are 24.5 dBi and 10 dBi, respectively, the noise level is 6dB, and the detection threshold is 100 dBm.

This localization experiments were performed on a lab machine using TensorFlow-GPU version 1.15. The lab machine has a NVIDIA GTX 1080 GPU, 6 GB of RAM, and 64-bit windows 10 operating system. The benchmark method is CNN-based mmWave localization [18]. Specifically, TFLearn is used to speed up the computations in the CNN model. In addition, we use the GPflow framework to implement the proposed DCGP model, where the Adam learning rate is set to 0.01, the convolution filter size is 3×3, and the number of inducing points is 384 at each layer.

3.2. Experimental Results

Fig. 3 presents the cumulative distribution functions (CDFs) of location errors achieved by DCGP and CNN. The 95th percentile errors for DCGP and CNN are 7.018 m and 14.289 m, respectively, while the median errors for DCGP and CNN are 2.559 m and 5.885 m, respectively. We can see that the proposed DCGP model achieves higher robustness than CNN. Using the same device with a 1080 GPU, the training time for DCGP and CNN are 533 minutes and 376 minutes, respectively. The higher training time of DCGP is due to more calculations for optimizing the kernel functions.

We also evaluate the impact of number of epochs on mean distance error. Fig. 4 shows the average distance errors over different numbers of epochs using CNN and DCGP, respectively. The range of number of epochs is from 50 to 550. It can be seen that for both CNN and DCGP models as the number of epochs is increased from 50 to 550, the distance error decreases. Specially, when the number of epochs is 50, the

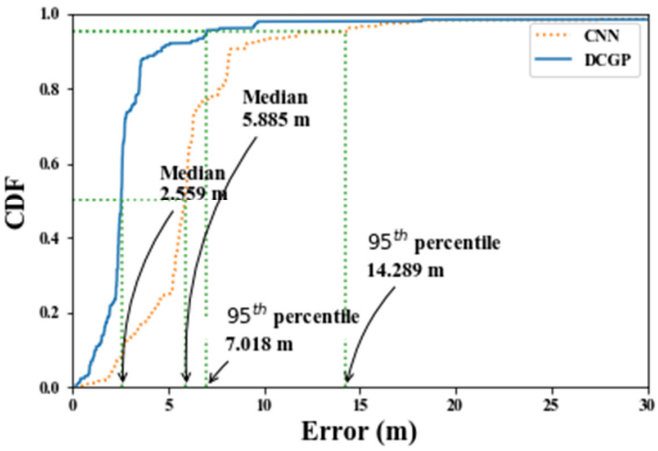

Fig. 3. CDFs of distance error achieved by the DCGP and CNN models.

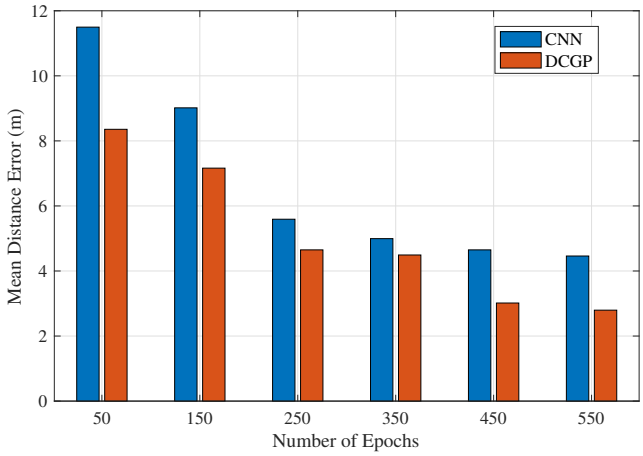

Fig. 4. Average distance errors over different numbers of epochs using CNN and DCGP.

mean distance errors for CNN and DCGP are 11.49 m and 8.35 m, respectively. At the 550th epoch, the mean distance errors of CNN and DCGP become 4.46 m and 2.79 m, respectively. The proposed DCGP model outperforms the CNN model for mmWave outdoor localization.

4. CONCLUSIONS

In this paper, we proposed a DCGP based approach for outdoor localization with mmWave beamforming image dataset. We presented the system architecture of mmWave based outdoor localization, including beamforming image construction and DCGP training, where the DCGP model is used to extract the features of mmWave beamforming images. Our experimental results demonstrated that the proposed DCGP method can achieve better outdoor localization accuracy using the 28 GHz mmWave beamforming image dataset, compared with the CNN-based baseline method.

5. REFERENCES

[1] José A del Peral-Rosado, Ronald Raulefs, José A López-Salcedo, and Gonzalo Seco-Granados, "Survey of cellular mobile radio localization methods: From 1G to 5G," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 1124–1148, Secondquarter 2018.

[2] Jait Purohit, Xuyu Wang, Shiwen Mao, Xiaoyan Sun, and Chao Yang, "Fingerprinting-based indoor and outdoor localization with LoRa and deep learning," in *Proc. IEEE GLOBECOM 2020*, Taipei, Taiwan, Dec. 2020, pp. 1–6.

[3] Avik Ray, Supratim Deb, and Pantelis Monogioudis, "Localization of LTE measurement records with missing information," in *Proc. IEEE INFOCOM 2016*, San Francisco, CA, Apr. 2016, pp. 1–9.

[4] Arash Shahmansoori, Gabriel E Garcia, Giuseppe Destino, Gonzalo Seco-Granados, and Henk Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, Mar. 2017.

[5] Zohair Abu-Shaban, Xiangyun Zhou, Thushara Abhayapala, Gonzalo Seco-Granados, and Henk Wymeersch, "Error bounds for uplink and downlink 3D localization in 5G millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 4939–4954, Aug. 2018.

[6] Ojas Kanhere, Shihao Ju, Yunchou Xing, and Theodore S Rappaport, "Map-assisted millimeter wave localization for accurate position location," in *Proc. IEEE GLOBECOM 2019*, Waikoloa, HI, Dec. 2019, pp. 1–6.

[7] Xuyu Wang, Lingjun Gao, and Shiwen Mao, "CSI-based fingerprinting for indoor localization: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 763–776, Jan. 2017.

[8] Xuyu Wang, Xiangyu Wang, and Shiwen Mao, "Deep convolutional neural networks for indoor localization with CSI images," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 1, pp. 316–327, Jan./Mar. 2020.

[9] X. Wang, X. Wang, and S. Mao, "Indoor fingerprinting with bimodal CSI tensors: A deep residual sharing learning approach," *IEEE Internet of Things Journal*, to appear, DOI: 10.1109/JIOT.2020.3026608.

[10] Zhiqing Wei, Yadong Zhao, Xinyi Liu, and Zhiyong Feng, "DoA-LF: A location fingerprint positioning algorithm with millimeter-wave," *IEEE Access J.*, vol. 5, pp. 22678–22688, Sept. 2017.

[11] João Gante, Gabriel Falcão, and Leonel Sousa, "Beamformed fingerprint learning for accurate millimeter wave positioning," in *Proc. IEEE VTC-Fall 2018*, Chicago, IL, Aug. 2018, pp. 1–5.

[12] João Gante, Gabriel Falcão, and Leonel Sousa, "Enhancing beamformed fingerprint outdoor positioning with hierarchical convolutional neural networks," in *Proc. IEEE ICASSP 2019*, Brighton, UK, May 2019, pp. 1473–1477.

[13] Toshiaki Koike-Akino, Pu Wang, Milutin Pajovic, Haijian Sun, and Philip V Orlik, "Fingerprinting-based indoor localization with commercial MMWave WiFi: A deep learning approach," *IEEE Access J.*, vol. 8, pp. 84879–84892, Apr. 2020.

[14] Guillermo Bielsa, Joan Palacios, Adrian Loch, Daniel Steinmetzer, Paolo Casari, and Joerg Widmer, "Indoor localization using commercial off-the-shelf 60 GHz access points," in *Proc. IEEE INFOCOM 2018*, Honolulu, HI, Apr. 2018, IEEE, pp. 2384–2392.

[15] Chenshu Wu, Feng Zhang, Beibei Wang, and KJ Ray Liu, "mmTrack: Passive multi-person localization using commodity millimeter wave radio," in *Proc. IEEE INFOCOM 2020*, Toronto, Canada, July 2020, pp. 1–9.

[16] Kenneth Blomqvist, Samuel Kaski, and Markus Heinonen, "Deep convolutional Gaussian processes," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Würzburg, Germany, Sept. 2019, pp. 582–597.

[17] Xiangyu Wang, Xuyu Wang, Shiwen Mao, Jian Zhang, Senthilkumar CG Periaswamy, and Justin Patton, "Indoor radio map construction and localization with deep gaussian processes," *IEEE Internet of Things Journal*, vol. 7, no. 11, pp. 11238–11249, 2020.

[18] João Gante, Gabriel Falcão, and Leonel Sousa, "Deep learning architectures for accurate millimeter wave positioning in 5G," *Neural Processing Letters*, vol. 51, no. 1, pp. 487–514, Aug. 2020.

[19] Michalis Titsias and Miguel Lázaro-Gredilla, "Doubly stochastic variational Bayes for non-conjugate inference," in *Proc. ICML 2014*, Beijing, China, June 2014, pp. 1971–1979.

[20] Yaniv Azar, George N Wong, Kevin Wang, Rimma Mayzus, Jocelyn K Schulz, Hang Zhao, Felix Gutierrez, DuckDong Hwang, and Theodore S Rappaport, "28 GHz propagation measurements for outdoor cellular communications using steerable beam antennas in New York City," in *Proc. IEEE ICC 2013*, Budapest, Hungary, June 2013, pp. 5143–5147.