Radar Aided 6G Beam Prediction: Deep Learning Algorithms and Real-World Demonstration

Umut Demirhan and Ahmed Alkhateeb
Arizona State University - Email: udemirhan, alkhateeb@asu.edu

***Abstract*—Adjusting the narrow beams at millimeter wave (mmWave) and terahertz (THz) MIMO communication systems is associated with high beam training overhead, which makes it hard for these systems to support highly-mobile applications. This overhead can potentially be reduced or eliminated if sufficient awareness about the transmitter/receiver locations and the surrounding environment is available. In this paper, efficient deep learning solutions that leverage radar sensory data are developed to guide the mmWave beam prediction and significantly reduce the beam training overhead. Our solutions integrate radar signal processing approaches to extract the relevant features for the learning models, and hence optimize their complexity and inference time. The proposed machine learning based radar-aided beam prediction solutions are evaluated using a large-scale real-world mmWave radar/communication dataset and their capabilities were demonstrated in a realistic vehicular communication scenario. In addition to completely eliminating the radar/communication calibration overhead, the proposed algorithms are able to achieve around 90% top-5 beam prediction accuracy while saving 93% of the beam training overhead. This highlights a promising direction for addressing the training overhead challenge in mmWave/THz communication systems.**

I. Introduction

Millimeter wave (mmWave) and terahertz (THz) communications systems rely on the beamforming gains of the narrow beams to achieve sufficient receive signal power. Finding the best narrow beam (or beam pair), however, requires high beam training overhead, which makes it hard for these systems to support highly mobile applications such as vehicular, drone, or augmented/virtual reality communications [1]. One important observation here is that the beam selection problem highly relies on the transmitter/receiver locations and the geometry/characteristics of the surrounding environment. This means that acquiring some awareness about the surrounding environment and the transmitter/receiver locations could potentially help the mmWave beam selection problem. An efficient way to acquire this awareness is by using the low-cost radar sensors such as those initially designed for automotive applications [2]. With this motivation, this paper investigates the potential of leveraging radar sensory data to guide the beam selection problem and provides the first machine learning based real-world demonstration for radar-aided beam prediction in a practical vehicular communication scenario.

Leveraging sensory data to guide the mmWave beam selection problem has gained increasing interest in the last few years [3]–[7]. In [3], [4], the authors proposed to leverage the sub-6GHz channels that are relatively easier to acquire to guide the beam selection problem. Acquiring sub-6GHz channels, however, still requires allocating wireless communication resources and probably additional control signaling. In [5], the position information of the user was leveraged by the base station to select the mmWave beam. The position information, though, may not be sufficient to accurately determine the best beam, which is also a function of the surrounding environment, especially in the non-line-of-sight scenarios. Further, acquiring accurate enough position information to adjust the narrow beams (i) may require expensive positioning systems at the user for the outdoor scenarios, and (ii) is hard to achieve for indoor communication. This motivated leveraging other data modalities for beam selection such as vision [6], which could be acquired at low-cost and without consuming any wireless communication/control resources, or radar data [7] which may operate at a different band than that used by the mmWave communication system. The proposed solution in [7], however, did not leverage machine learning and relied only on classical calibration techniques for the radar and communication systems, which could be expensive and hard to implement in reality. Further, the results in [7] relied only on simple computer simulations and are hard to scale to real-world scenarios and practical hardware.

In this work, we develop machine learning based algorithms for radar-aided mmWave beam prediction and demonstrate their performance using a real-world dataset in a realistic vehicular communication scenario. The main contributions of the paper can be summarized as follows: (i) We formulate the radar-aided beam prediction problem considering practical radar and communication models, (ii) we then develop efficient machine learning algorithms that leverage classical signal preprocessing approaches for extracting the relevant features such as range-velocity, range-angle, and range-angle-velocity maps, (iii) leveraging a large-scale dataset that is collected using a 77 GHz frequency-modulated continuous wave (FMCW) radar and communication mmWave phased arrays, we evaluate and demonstrate the performance of the proposed radar-aided beam prediction approaches in a realistic vehicular communication scenario and draw some insights about the various algorithms in terms of prediction accuracy, processing time, inference latency, and complexity overhead.

II. System Model

The considered system in this paper consists of a base station and a mobile user. The base station employs two main components: (i) A mmWave communication terminal equipped with a phased array that is used to communicate with

the mobile user, and (ii) an FMCW radar that is leveraged to aid the selection of the mmWave communication beam. The system model is illustrated in Fig. 1. In the next two subsections, we briefly describe the system and signal models of the communication and radar components.

A. Radar Model

In our system, the base station adopts an FMCW radar. The objective of this radar is to provide observations of the environment. The FMCW radar achieves this objective by transmitting chirp signals whose frequency changes continuously with time. More formally, the FMCW radar transmits a linear chirp signal starting at an initial frequency f_c and linearly ramping up to $f_c + \mu t$, given by

$$s^{\text{tx}}_{\text{chirp}}(t) = \begin{cases} \sin(2\pi[f_c\, t + \frac{\mu}{2}\, t^2]) & \text{if } 0 \leq t \leq T_c \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $\mu = B/T_c$ is the slope of the linear chirp signal with B and T_c representing the bandwidth and duration of the chirp.

A single radar measurement is obtained from the frame of duration T_f. In each frame, L chirp waves are transmitted with T_s waiting time between them. After the transmission of the last chirp, no other signals are transmitted until the completion of the frame. Mathematically, we can write the transmitted signal of the radar frame as

$$s^{\text{tx}}_{\text{frame}}(t) = \sqrt{\mathcal{E}_t} \sum_{l=0}^{L-1} s^{\text{tx}}_{\text{chirp}}(t-(T_c+T_s)\cdot l), \quad 0 \leq t \leq T_f \quad (2)$$

where $\sqrt{\mathcal{E}_t}$ is the transmitter gain. The given transmitted signal is reflected from the objects in the environment, and received back at the radar.

At the receiver, the signal obtained from an antenna is passed through a quadrature mixer that combines the transmit with receive signals resulting in the in-phase and quadrature samples. After that, a low-pass filter is applied to the mixed signals. The resulting signal, referred to as intermediate frequency (IF) signal, reflects the frequency and phase difference between the transmit and receive signals. If a single object exists in the environment, then the receive IF signal of a single chirp can be written as

$$s^{\text{rx}}_{\text{chirp}}(t) = \sqrt{\mathcal{E}_t \mathcal{E}_r} \exp\left(j2\pi[\mu\tau t + f_c\tau - \frac{\mu}{2}\tau^2]\right), \quad (3)$$

where $\sqrt{\mathcal{E}_r}$ is the channel gain of the object which depends on the radar cross section (RCS) and the path-loss, $\tau = 2d/c$ is the round-trip delay of the reflected signal through the object with d denoting the distance between the object and the radar, and c representing the speed of light.

The receive IF signal, $s^{\text{rx}}_{\text{chirp}}(t)$ is then sampled at the sampling rate of the analog to digital converter (ADC), f_s, producing S samples for each chirp. Finally, the ADC samples from each frame are collected. For an FMCW radar with M_r receive antennas, each having the described RF receive chain, the resulting measurements (raw-data) of one frame can be denoted by $\mathbf{X} \in \mathbb{C}^{M_r \times S \times L}$. In the following subsection, we describe the communication model.

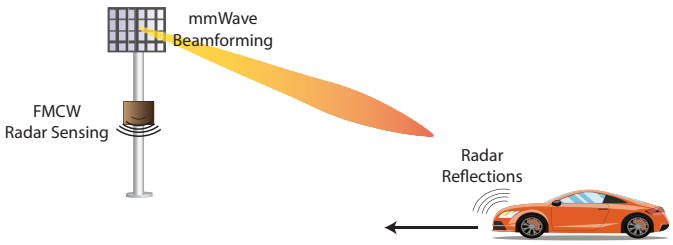

Fig. 1. The system model where the FMCW radar information at the basestation is leveraged to select the beam that serves the mobile user.

B. Communication Model

The considered base station employs a mmWave transceiver with M_c antennas and use it to communicate with a single-antenna mobile user. Adopting a narrowband channel model, the channel between the user and the base station can be expressed as

$$\mathbf{h} = \sum_{p=0}^{P-1} \alpha_p \mathbf{a}(\phi_p, \theta_p) \quad (4)$$

where α_p denotes the complex gain, ϕ_p and θ_p represent transmit azimuth and elevation angles of the p-th path at the base station, and $\mathbf{a}(\phi_p, \theta_p)$ is the array response vector of these angles. In the downlink, the base station transmits the data symbol s_d to the user via the beamforming vector $\mathbf{f} \in \mathbb{C}^{M_c}$. The receive signal at the user can be written as

$$y = \sqrt{\mathcal{E}_c}\mathbf{h}^H \mathbf{f} s_d + n \quad (5)$$

where $n \sim \mathcal{CN}(0, \sigma^2)$ is the additive white Gaussian noise and $\sqrt{\mathcal{E}_c}$ is the transmitter gain of the basestation. For the selection of the beamforming vectors, we define the beamforming codebook of N vectors by $\boldsymbol{\mathcal{F}}$, where the n-th beamforming vector is denoted by $\mathbf{f}_n \in \boldsymbol{\mathcal{F}}$, $\forall n \in \{0, \ldots, N-1\}$. Hence, $\mathbf{f}$ in (5) is restricted to the beams in the codebook. With this model, the index of the optimal beam, $n^\star$, can be obtained by the SNR maximization problem, i.e.,

$$n^\star = \underset{n}{\text{argmax}} \ |\mathbf{g}^H \mathbf{f}_n|^2 \quad \text{s.t.} \quad \mathbf{f}_n \in \boldsymbol{\mathcal{F}} \quad (6)$$

where the optimal solution can be obtained by an exhaustive search over the possible beamforming vectors.

III. Machine Learning and Radar Aided Beamforming

In this section, we formally define the radar based beamforming problem, building upon the described system model in Section II. Then, we present the proposed solution.

A. Problem Definition

In this paper, we seek to leverage the radar measurements $\mathbf{X}$ in determining the optimal communication beamforming vector $\mathbf{f}_{n^\star}$. First, let us introduce the subscript k to indicate the k-th radar frame. The radar measurements during this frame will then be denoted as $\mathbf{X}_k$. Further, we add this subscript k to the beamforming index and the beamforming vector used in this k-th frame, to be n_k, $\mathbf{f}_{n_k}$. If a single-user exists in

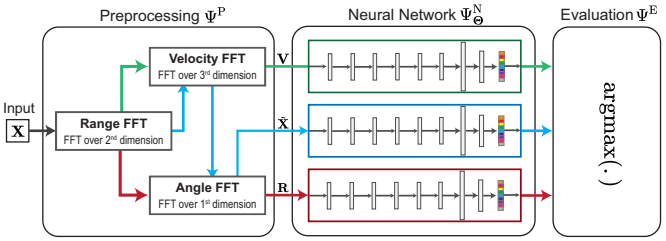

Fig. 2. The figure illustrates the radar processing procedures for the three proposed approaches. The results of these processing procedures, $\tilde{\mathbf{X}}, \mathbf{R}, \mathbf{V}$ are the inputs to the neural networks.

the line-of-sight (LOS) of the base station, then the radar measurements could potentially include useful information $\mathbf{X}_k$ about its position/orientation with respect to the base station. This position/orientation information could be leveraged to guide the optimal beam selection. To formulate that, we define the mapping function $\Psi_{\boldsymbol{\Theta}}$ to capture the mapping from the radar observations to the optimal beamforming index, given by

$$\Psi_{\boldsymbol{\Theta}} : \{\mathbf{X}_k\} \rightarrow \{n^\star_k\} \tag{7}$$

Our objective is then to design the mapping function $\Psi_{\boldsymbol{\Theta}}$ to be able to map the radar measurements to the optimal beam index $n^\star$. Towards this objective, we investigate the possible designs of the mapping function, and learn the set of parameters $\boldsymbol{\Theta}$. Mathematically, we can express this objective by the following optimization problem that aims at finding the mapping function and the optimal set of parameters $\boldsymbol{\Theta}$, that maximizes the accuracy in predicting the optimal beam:

$$\Psi^\star_{\boldsymbol{\Theta}^\star} = \arg\max_{\Psi_\Theta} \frac{1}{K} \sum_{k=1}^{K} \mathbf{1}_{\{n^*_k = \Psi_{\boldsymbol{\Theta}}(\mathbf{X}_k)\}} \tag{8}$$

where $\mathbf{1}_E$ is the indicator function of the event E, i.e., $\mathbf{1}_E = 1$ if E occurs, and $\mathbf{1}_E = 0$ otherwise. Next, we present our machine learning aided solution.

B. Proposed Framework

In this section, we present our deep learning and radar aided beam prediction approach. Our solution integrates radar pre-processing and deep neural networks. This targets reducing the complexity of the learning task and enables efficient training with reasonable dataset sizes. To formalize our approach, we first decompose the radar to beam mapping function into three components: (i) The pre-processing function $\Psi^{\mathrm{P}}(.)$, (ii) the neural network function of the parameters $\boldsymbol{\Theta}$, $\Psi^{\mathrm{N}}_{\boldsymbol{\Theta}}(.)$, and (iii) the evaluation function $\Psi^{\mathrm{E}}(.)$. Then, we can write the radar to beam mapping function as $\Psi_{\boldsymbol{\Theta}}(\mathbf{X}) = \Psi^{\mathrm{E}}(\Psi^{\mathrm{N}}_{\boldsymbol{\Theta}}(\Psi^{\mathrm{P}}(\mathbf{X})))$. With the decomposition, we can define our solution in terms of the pre-processing, neural networks, and evaluation functions. In the following, we present our approach via the subsections of each function. First, we describe the proposed pre-processing approach.

1) Preprocessing: In pricipal, given the radar raw measurements $\mathbf{X} \in \mathbb{C}^{M_r \times S \times L}$, three important features that could be extracted are the range, the angles, and the velocity of the moving objects in the environment. Based on that, we propose three different pre-processing approaches, as illustrated in Fig. 2. Each approach leverages a certain set of these quantities. To mathematically define the pre-processing approaches, let us denote the 2D and 3D Fourier transforms by $\mathcal{F}_{\mathrm{2D}(.)}$ and $\mathcal{F}_{\mathrm{3D}}(.)$, respectively. Next, we describe the three pre-processing approaches.

Range-Angle Maps: The first approach aims at utilizing the range and angle information. For this purpose, first, with an FFT in the direction of the time samples, referred to as the Range FFT, we obtain the chirp signal in the frequency domain. With an another FFT in the direction of the receive antenna samples, referred to as the Angle FFT, the angular information can be obtained. An FFT of a larger size, M_F, can be applied with zero-padding to over-sample the angles. Finally, we can construct the final range-angle map by combining resulting range-angle information for each chirp sample, i.e., $\mathbf{R} = \Psi^{\mathrm{P}}_{\mathrm{A}}(\mathbf{X}) = \sum_{l=1}^{L} |\mathcal{F}_{\mathrm{2D}}(\mathbf{X}_{:,:,l})|$.

Range-Velocity Maps: Alternatively, we consider the range-velocity maps. To construct these maps from the radar measurements, two FFTs through the time samples and chirp samples are applied. Similarly to the previous approach, first, the Range FFT is utilized. Differently, the second FFT is applied through the chirp samples, referred to as the Velocity FFT. It simply returns the phase shift over the consecutive chirp samples. This phase shift is caused by the Doppler shift, and it contains the velocity information. Finally, again by combining the range-velocity information of the different receive antenna samples, we obtain the final range-velocity map by $\mathbf{V} = \Psi^{\mathrm{P}}_{\mathrm{V}}(\mathbf{X}) = \sum_{m=1}^{M} |\mathcal{F}_{\mathrm{2D}}(\mathbf{X}_{m,:,:})|$.

Radar Cube: The previous approaches combine the angle or velocity dimensions, reducing the information to a 2D map. Without a dimensionality reduction, we apply the range, velocity, and angle FTTs, and obtain the radar cube. The resulting radar cube contains all the information of the range, velocity, and angle of the targets. It can be considered as the stack of range-angle maps of each velocity value. The operation can be mathematically described as $\tilde{\mathbf{X}} = \Psi^{\mathrm{P}}_{\mathrm{C}}(\mathbf{X}) = |\mathcal{F}_{\mathrm{3D}}(\mathbf{X})|$.

After the alternative modalities of the radar information are extracted, the data is fed into the neural networks. The described radar processing approaches bring different pre-processing complexity and input size. In particular, while the radar cube requires a 3D FFT presenting the most detailed information, it suffers from the high number of dimensions. In contrast, the range-angle and range-velocity images only require 2D FFTs and provide smaller input sizes. The further evaluation of the complexity is carried out in Section V. In Table I, we summarize the pre-processing complexity and size of each data entry. Next, we present the deep neural networks

TABLE I
COMPLEXITY AND MEMORY REQUIREMENTS OF DIFFERENT INPUTS

Network Input	Preprocessing Complexity	Input Size
$\tilde{\mathbf{X}}$	$\mathcal{O}(M_r SL(\log S + \log L + \log M_r))$	$M_r SL$
$\mathbf{R}$	$\mathcal{O}(M_r SL \log S + M_F SL \log M_F)$	$M_F S$
$\mathbf{V}$	$\mathcal{O}(M_r SL(\log S + \log L))$	SL

TABLE II
DEEP NEURAL NETWORK ARCHITECTURES FOR DIFFERENT INPUT TYPES

NN Layers	Radar Cube ($\tilde{\mathbf{X}}$)	Range-Velocity (V)	Range-Angle-64 (R)	Range-Angle-4 (R)
Input	$4 \times 256 \times 128$	$1 \times 256 \times 128$	$1 \times 256 \times 64$	$1 \times 256 \times 4$
CNN-1	Output Channels: 8, Kernel: (3, 3), Activation: ReLU			
CNN-2	Output Channels: 16, Kernel: (3, 3), Activation: ReLU			
AvgPool-1	Kernel: (2, 1)		N/A	
CNN-3	Output Channels: 8, Kernel: (3, 3), Activation: ReLU			
AvgPool-2	Kernel: (2, 2)			Kernel: (2, 1)
CNN-4	Output Channels: 4, Kernel: (3, 3), Activation: ReLU			
AvgPool-3	Kernel: (2, 2)			Kernel: (2, 1)
CNN-5	Output Channels: 2, Kernel: (3, 3), Activation: ReLU			
AvgPool-4	Kernel: (2, 2)			Kernel: (2, 1)
FC-1	Input Size: 512, Output Size: 256, Activation: ReLU			
FC-2	Input Size: 256, Output Size: 128, Activation: ReLU			
FC-3	Input Size: 128, Output Size: 64			

adopted for each modality of the data.

2) Neural Network Modeling: For the neural networks, to keep the complexity of the approach low, we rely on a comparably simple deep learning model with a design with convolutional and fully-connected (FC) layers. Specifically, the deep neural networks DNNs comprise 8 total layers. The first five layers are the convolutional layers with the rectified linear unit (ReLU) activation functions. In addition, the average pooling is applied after the activation of the convolutional layers to decrease the size of the data. Finally, the output of the fifth convolutional layer is connected to a set of three FC layers, providing N outputs. The each entry of the output indicates a beam.

As the proposed inputs of the neural networks are of different size and dimensions, the same network cannot be applied to the all types of the inputs. Therefore, for different modalities of the radar data, the input, output and kernel size of the DNN layers are adjusted to keep the network size reasonable and similar while providing a comparably good performance. Specifically, we adjust the networks for our dataset, which will be described in Section IV. In this dataset, the system parameters are given by $S = 256$, $L = 128$, $M_r = 4$ and $N = 64$. With these parameters and $M_F \in \{4, 64\}$, the designed DNN architectures are summarized in Table II.

Neural Network Objective: To train the neural networks with the aim of finding the optimal parameters $\boldsymbol{\Theta}^\star$, we can write the following optimization problem that aims at minimizing the loss between the output of the network and the optimal beam values, $n_k^\star \in \{0, \ldots, N-1\}$:

$$\boldsymbol{\Theta}^\star = \arg\min_{\boldsymbol{\Theta}} \frac{1}{K} \sum_{k=1}^{K} \mathcal{L}\left(\Psi_{\boldsymbol{\Theta}}^{\mathrm{N}}\left(\Psi^{\mathrm{P}}(\mathbf{X}_k)\right), n_k^\star\right) \tag{9}$$

where $\mathcal{L}(.,.)$ denotes the loss function. As our problem is a multi-class classification problem, we utilize the cross-entropy loss given by $\mathcal{L}(\hat{\mathbf{b}}, \mathbf{b}) = -\frac{1}{N}\sum_{n=0}^{N-1} b_n \log(\hat{b}_n)$ where $\mathbf{b} = [b_0, \ldots, b_{N-1}]$ is the one-hot encoded vector of the optimal beam $n_k^\star$ and $b_n = \Psi_{\boldsymbol{\Theta}}^{\mathrm{N}}(\Psi^{\mathrm{P}}(\mathbf{X}_k))$ is the output of the neural network.

3) Evaluation: To evaluate the output of the neural network in terms of the objective function in (8), we need to select a single beam from the soft output of the neural network.

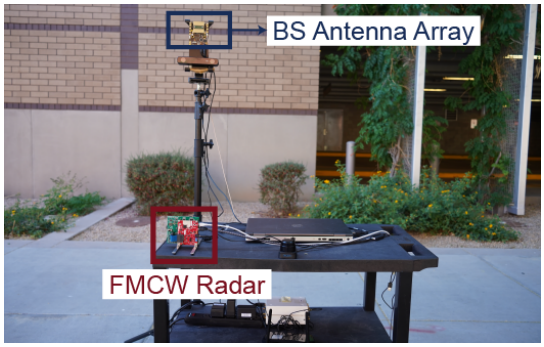

Fig. 3. Data collection system setup equipped with a TI AWR2243 mmWave FMCW radar working at 77 GHz operating frequency and a mmWave MIMO antenna array operating at 60 GHz.

For this purpose, the maximum of the neural network output can be selected as the prediction of the optimal beamforming vector. This can be mathematically provided by setting $\Psi^{\mathrm{E}}(.) = \arg\max(.)$, completing our solution. In the following section, we describe our dataset adopted in the training and evaluation of the solutions.

IV. REAL-WORLD DATASET

To accurately evaluate the performance of the proposed machine learning based radar-aided beam prediction approaches in a realistic environment, we built a real-world dataset with radar and wireless measurements. In this section, we describe our testbed and present the dataset collection scenario.

A. Data Collection Testbed

We adopt the Testbed-1 of the DeepSense 6G dataset [8]. The testbed comprises two units: The stationary unit (Unit 1), and the mobile unit (Unit 2). Among other sensors, unit 1 employs an FMCW radar (AWR2243BOOST) which has 3 transmit and 4 receive antennas, and a mmWave receiver at 60 GHz which adopts a uniform linear array (ULA) with $M_c = 16$ elements. The unit 2 utilizes a 60 GHz quasi-omni antenna, acts as a transmitter and is always oriented towards the receiver antenna of unit 1.

The phased array of unit 1 utilizes an over-sampled beamforming codebook of $N = 64$ vectors, which are designed to cover the field of view. It captures the receive power by applying the beamforming codebook elements as a combiner. The combiner providing the most power is taken as the optimal beamforming vector. For the radar, we only activated one of the TX antennas, while the data from $M_r = 4$ RX antennas were captured. We adopted a set of radar parameters given by $B = 750$ MHz, $\mu = 15$ MHz/us, $L = 128$ chirps/frame, $S = 256$ samples/chirp. These settings provide the maximum range of 45m and the maximum velocity of 56 km/h, which are well-fit for the scenario illustrated in Fig. 4. Next, we present the dataset and collection scenario.

B. Development Dataset

For the evaluation, we used the testbed described in Section IV-A and adopted Scenario 9 of the DeepSense 6G dataset

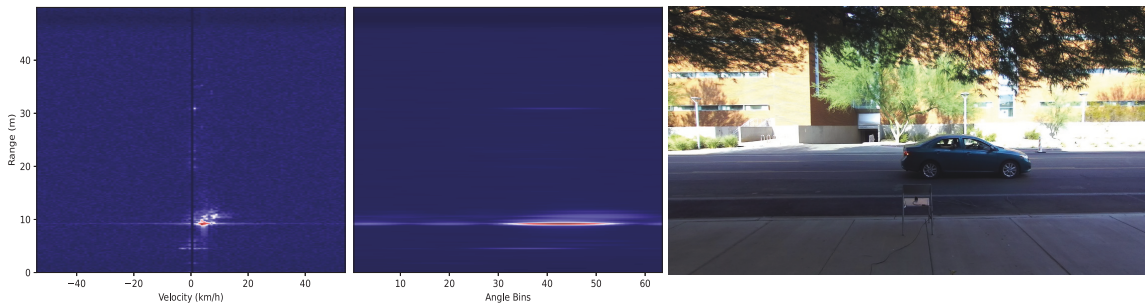

Fig. 4. A sample from the dataset is shown with the current environment image (right) and the corresponding range-angle (middle) and range-velocity (left) images. The car is on the right part of the camera image, moving away from the vertical angle of the radar device. The range-angle image shows the position at approximately 9m distance on the right hand side, while the range-velocity image indicates the increasing relative velocity and range.

[8]. In this scenario, a passenger at the back seat of the car holds the transmitter. As shown in Fig. 4, the car passes by the stationary unit (Unit 1) which collects the radar and beam training measurements. During the data collection, the road was actively used by the other cars, pedestrians and bikers. Our testbed collected and saved the radar measurements and the received power at each communication beam.

In the construction of the dataset, the beam providing the highest power is saved as the optimal beamforming vector. The data is cleaned by only keeping the samples with the target car in sight. This cleaning operation is performed manually through the inspection of the RGB images that are captured from a camera attached next to the antenna array. The data samples with the other elements (cars, pedestrians and bikers) are also kept to reflect the realistic environment. The final dataset comprises 6319 samples, which are separated with a $70/20/10\%$ split for the training, validation and testing. A sample from the dataset through the extracted range-angle and range-velocity images are shown in Fig. 4.

V. Results

In this section, we evaluate and compare the performance of the proposed solutions. In particular, we compare the DNN based solutions adopting the radar cube, range-velocity and range-angle maps, respectively, and a simple baseline algorithm. Adopting our dataset described in Section IV, the solutions are compared in terms of their prediction accuracy, complexity/inference time, and the required dataset sizes.

Baseline Algorithm: As the baseline algorithm, we adopt a look-up table mapping the given position of the maximum point in the range-angle image to the most-likely beams. For the top-K accuracy, we select K different beams corresponding to the largest points in the range-angle map.

Training and Evaluation: For the evaluation of the neural networks, we trained the DNN models summarized in Table II using the Adam algorithm [9] with a learning rate 0.001, batch size 32, and a decay factor $\gamma = 0.1$ which is applied after every 10 epochs. The networks are trained for 40 epochs and the network parameters showing the best top-1 accuracy over the validation dataset is saved for evaluation. The network training operation is carried out for 5 separate instances, and the average performance is shown in the following results. Next, we compare the beam prediction accuracy of the solutions.

Beam Prediction Accuracy: As shown in Fig. 5, the range-angle map based deep learning solutions over-perform the range-velocity and radar cube solutions. In comparison to the deep learning solutions, the baseline solutions show an inferior performance. Specifically, the baseline with 64-point angle FFT provides 33% accuracy, while the deep learning models provide at least 8% better top-1 accuracy. This shows the robustness and applicability of the deep learning models to the real-world data. The radar cubes contain the range-angle maps of different velocity values, however, it cannot perform similarly to the 4-point range-angle solution. This is potentially due to the large size and complexity of the input and comparable simplicity of the deep learning model.

For top-5 accuracy, the baseline solution only reaches up to 63% and has a smaller gain when increasing the K values. In comparison, the top-3 and top-5 accuracy of the range-angle images reach up to 79.7% and 93.5%, outperforming the baseline solution by a large margin. The other deep learning based approaches show improvements similar to the range-angle based deep learning solutions with increasing K values. Based on the presented results, we conclude that the deep learning solutions show clear potential for real applications, especially with the top-3 and top-5 results. Moreover, the comparison of the range-angle solutions with different angle FFT sizes show the advantage of the maps generated with higher resolution, which shows the advantages of generating maps with higher resolution for deep learning solutions.

Complexity: In Fig. 6, we compare the complexity of the DNNs in terms of the number of parameters, preprocessing and network inference durations. First, in the design of the DNN models, we aimed to keep the number of parameters similar. To that end, the radar cube, range-velocity and range-angle based solutions are comprised of approximately 175k parameters. The baseline solutions only require 1024 and 16384 parameters for 4- and 64-point angle FFTs. These values correspond to each pixel in the range-angle maps, each of which is utilized to represent the most likely beam for each point. Second, we compare the inference duration of the neural networks as shown on the left part of Fig. 6. The deep learning solutions show similar inference durations since the

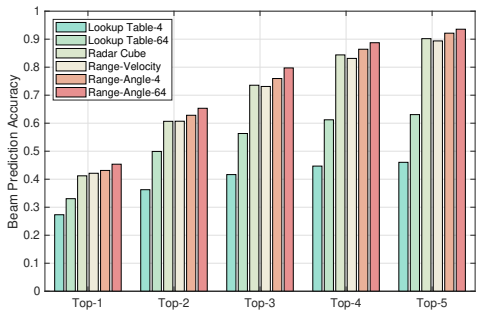

Fig. 5. The top-k test accuracy of the proposed approaches. Range-angle map based deep learning solutions outperform the other deep learning solutions.

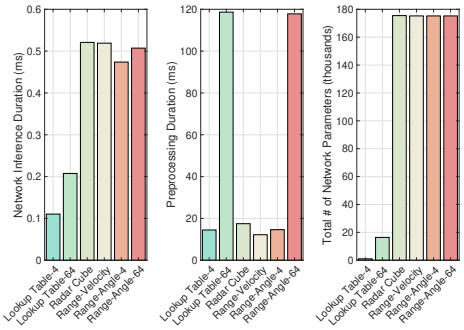

Fig. 6. The complexity of the proposed radar-aided beam prediction approaches are compared in terms of the network inference time, preprocessing duration, and number of (neural network) parameters.

number of parameters and network architectures are designed similarly. Third, the middle figure in Fig. 6 illustrates the radar preprocessing durations, where the larger angle FFT adopted in the range-angle maps causes a significant additional time. This present a trade-off between the beam prediction accuracy and complexity of the solution. Depending on the hardware availability, one may prefer to design a solution with higher resolution maps and better accuracy. Without an oversampling of the angle FFT dimension, all the approaches show similar durations, presenting an advantage for the better performing solutions.

Impact of Dataset Size: In Fig. 7, we show the average accuracy of the trained networks on the same test samples, while only a subset of the training samples are utilized. The lines of the same input for different K values show similar behavior with different scaling and accuracy levels. In the figure, the performance of the 64-point range-angle solution requires $10-20\%$ of the data for starting to saturate with a better accuracy, while the other solutions need around $20-30\%$ of the training data. This might be due to the easier interpretability of the high-resolution range-angle maps for the beam prediction task, potentially requiring less learning and transformation. The radar-cube starts to saturate particularly late, and may slightly increase with more data, indication a potential benefit from a larger dataset or more complex models. Nevertheless, the range angle/velocity based solutions generalize the problem well and can perform well with smaller

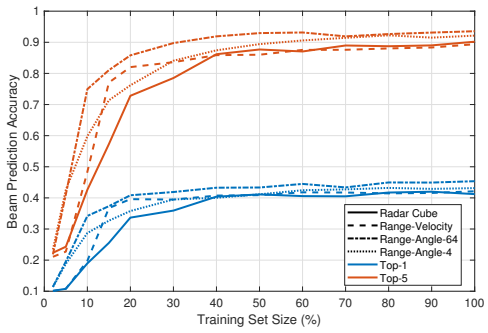

Fig. 7. Top-1 and Top-5 beam prediction accuracies of the proposed radar-aided beam prediction approaches when a fixed percentage of the training dataset is utilized.

datasets.

VI. Conclusions

In this paper, we developed machine learning based radar-aided mmWave beam prediction approaches and evaluated their performance using a real-world dataset. In a practical vehicular communication scenario and using a low-cost radar, we demonstrated that capability to efficiently predict the mmWave communication beam with reasonable pre-processing/inference latency. In particular, out of a 64-beam codebook, our approaches achieves around 90% top-5 beam prediction accuracy while requiring only a few milliseconds in pre-processing the radar data and inferring the best beams. This highlights a promising approach for enabling highly-mobile mmWave/THz communication applications.

VII. Acknowledgement

This work is supported by the National Science Foundation under Grant No. 2048021.

References

[1] A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic, "Deep learning coordinated beamforming for highly-mobile millimeter wave systems," *IEEE Access*, vol. 6, pp. 37 328–37 348, 2018.

[2] B. Ginsburg, et al, "A multimode 76-to-81GHz automotive radar transceiver with autonomous monitoring," in *2018 IEEE Int. Solid-State Circuits Conf. (ISSCC)*, 2018, pp. 158–160.

[3] M. Alrabeiah and A. Alkhateeb, "Deep learning for mmWave beam and blockage prediction using sub-6GHz channels," *IEEE Trans. Commun.*, pp. 1–1, 2020.

[4] A. Ali, N. Gonzalez-Prelcic, and R. W. Heath, "Millimeter wave beam-selection using out-of-band spatial information," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1038–1052, Feb 2018.

[5] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, "Position-aided millimeter wave V2I beam alignment: A learning-to-rank approach," in *2017 IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, 2017, pp. 1–5.

[6] M. Alrabeiah, A. Hredzak, and A. Alkhateeb, "Millimeter wave base stations with cameras: Vision-aided beam and blockage prediction," in *2020 IEEE 91st Veh. Technol. Conf (VTC2020-Spring)*, 2020, pp. 1–5.

[7] A. Ali, N. González-Prelcic, and A. Ghosh, "Passive radar at the roadside unit to configure millimeter wave vehicle-to-infrastructure links," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14 903–14 917, 2020.

[8] A. Alkhateeb, G. Charan, T. Osman, A. Hredzak, and N. Srinivas, "DeepSense 6G: A large-scale real-world multi-modal sensing and communication dataset," *available on arXiv*, 2021. [Online]. Available: https://www.DeepSense6G.net

[9] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.