

Enabling ISAC in Real World: Beam-Based User Identification with Machine Learning

Umut Demirhan and Ahmed Alkhateeb

Abstract—Leveraging perception from radar data can assist multiple communication tasks, especially in highly-mobile and large-scale MIMO systems. One particular challenge, however, is how to distinguish the communication user (object) from the other mobile objects in the sensing scene. This paper formulates this *user identification* problem and develops two solutions, a baseline model-based solution that maps the angle of the object from the radar scene to communication beams and a scalable deep learning solution that is agnostic to the number of candidate objects. Using the DeepSense 6G dataset, which has real-world measurements, the developed deep learning approach achieves more than 89% communication user identification accuracy on the test set, highlighting a promising path for enabling integrated radar-communication applications in the real world.

I. INTRODUCTION

Employing large antenna arrays in millimeter wave (mmWave) communication systems is essential to provide sufficient beamforming gains and receive power. Minimizing the beam training overhead, however, is a challenging task especially in highly-mobile applications. Towards addressing this challenge, integrated sensing and communications (ISAC) can be a keystone, where wireless communications are aided by radar sensing information. This information can be utilized to build perception of the environment and the target objects that affect the communication channels. One main challenge in this framework, however, is that the objects determined in both the sensing and communication domains need to be matched to facilitate accurate sensing aid to communication in scenarios with multiple mobile objects in the sensing scenes. We term this problem *user identification*. In this paper, we investigate how to solve the user identification problem in the radar sensing scenes via the use of selected beam information and evaluate the feasibility using a real-world demonstration.

Various types of sensing information, e.g., camera images [1], and radar [2]–[5], have been considered for aiding communication objectives. In [1], images captured by a camera attached to the basestation are considered for the beam and blockage prediction. Similarly, the measurements from an out-of-band radar have been studied for beam prediction [2], [3] and blockage prediction [4]. To realize radar-aided systems in the real world, however, there are many major questions to be answered. To this end, although [3], [4] included an evaluation in a real-world setup, they were limited to the single-target scenarios, and the user identification has not been considered.

In the joint sensing and communication (ISAC) literature, particularly where sensing and communication antennas are

co-located, prior work has explored user identification [5]–[7]. These approaches typically rely on metrics like Euclidean distance [5] and Kullback-Leibler divergence [7] applied to radar-estimated locations with the initial location from the communication-based identification, which necessitates estimating these locations through communication links. In [6], a distributed vehicular setup leverages GPS position knowledge of communication users. However, this body of work diverges significantly from our approach in several key aspects. Specifically, it has not: (i) considered an off-the-shelf external radar, which presents unique challenges and opportunities; (ii) evaluated performance in a real-world system, limiting its practical applicability; (iii) utilized machine learning, a crucial component for future ISAC systems as highlighted in [8]; and (iv) employed only the readily available communication beam index for identification, simplifying deployment and reducing complexity. Our work addresses these gaps, offering a novel and practical approach for ISAC user identification.

Therefore, in this work, we address the user identification problem in radar-aided communication systems employing out-of-band frequency-modulated continuous wave (FMCW) radars. Specifically, we leverage radar-generated measurements of detected objects and the communication beam indices selected to serve each user, without requiring feedback or additional state estimation. To solve this problem, we develop a robust and scalable deep neural network (DNN) solution. Crucially, the developed model is agnostic to the number of detected objects, enabling it to operate effectively regardless of object density. For evaluation, we constructed a real-world dataset as part of the DeepSense 6G framework [9] and benchmarked our approach against baseline solutions. Our evaluations demonstrate that the proposed DNN solution achieves 89.3% accuracy on the test set, representing a gain of over 15 – 25% compared to the baselines, and showcasing its promising potential for real-world radar-aided communication systems.

II. SYSTEM MODEL

For the system model, as illustrated in Fig. 1, we consider a single mobile user served by a basestation, along with multiple candidate targets. The mobile user carries a single antenna mmWave receiver. Meanwhile, the basestation is equipped with (i) a mmWave antenna array that is used to communicate with the mobile user and (ii) an off-the-shelf FMCW radar that is leveraged to aid the mmWave communication functions.

A. Communication Model

For communications, we consider a MISO channel with N element antenna array at the basestation. For simplicity, in

The authors are with the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ, 85281 USA (Email: udemirhan, alkhateeb@asu.edu). This work was supported by the National Science Foundation (NSF) under Grant No. 2048021.

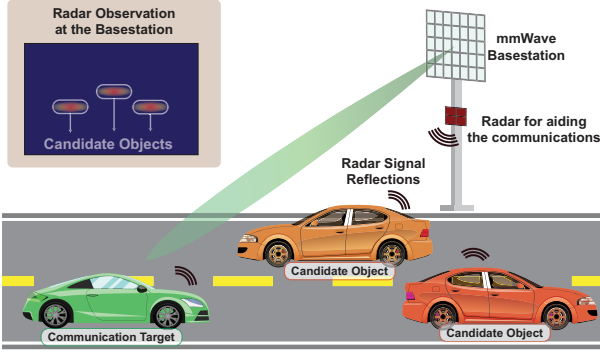


Fig. 1. The system model, where a radar-equipped mmWave basestation communicates with a user, while radar data to aid the communications is collected from all the available targets in the environment.

the following, we assume that the transmitter antennas form a uniform linear array (ULA). Nevertheless, this assumption can easily be relaxed to accommodate other antenna formation. Adopting a geometric channel model with P paths, we can write [3]

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

where α_p and θ_p denote the complex coefficient and azimuth angles of the p -th path. The function $\mathbf{a}(\theta)$ is defined as the array response vector of the basestation antenna array in the direction of θ . At the downlink, the basestation applies a beamforming vector $\mathbf{f} \in \mathbb{C}^N$ to the information s and transmits this signal over the channel. Hence, the signal received by the user can be written as

$$y = \sqrt{\rho} \mathbf{h}^H \mathbf{f} s + n, \quad (2)$$

with ρ being the transmitter power gain of the basestation and $n \sim \mathcal{N}(0, \sigma^2)$ being the additive white Gaussian noise. The beamforming vector \mathbf{f} is selected from a codebook \mathcal{F} of B beams. The b -th beamforming vector of the codebook is denoted by \mathbf{f}_b . With this model, the index of the optimal beam can be obtained by the beamforming gain maximization problem given as

$$b^* = \arg \max_{b \in \{1, \dots, B\}} |\mathbf{h}^H \mathbf{f}_b|^2, \quad (3)$$

where the optimal solution can be obtained by an exhaustive search over the beams of the codebook.

B. Radar Model

As described, our system model adopts an FMCW radar at the basestation. This radar aims to provide measurements to aid the communication system. Specifically, the radar transmits a sequence of linear up-chirps, given as [10]

$$s_{\text{chirp}}^{\text{tx}}(t) = \sin(2\pi[f_c t + \frac{\mu}{2} t^2]), \quad (4)$$

for $t \in [0, T_c]$, and 0 otherwise. The chirp signal starts from the frequency f_c and goes up to $f_c + \mu T_c$ with a constant slope $\mu = B/T_c$ within the duration of the signal. B and T_c represent the bandwidth and duration of a chirp signal.

The transmission of a sequence of chirps comprises a radar frame. Specifically, a sequence of M_c chirps, with T_w waiting

time between two consecutive chirps, is transmitted in each radar frame. With the transmission of a radar frame, the signal is reflected from the objects in the environment and received back at the receiver. The FMCW radar receiver mixes the current transmit signal with the received signal. The resulting signal is called an intermediate frequency (IF) signal. If there is a single object in the environment at a distance d , the IF signal from a single chirp can be written as

$$s_{\text{chirp}}^{\text{rx}}(t) = \sqrt{\mathcal{E}_t \mathcal{E}_r} \exp \left(j2\pi \left[\mu \tau_{rt} t + f_c \tau_{rt} - \frac{\mu}{2} \tau_{rt}^2 \right] \right), \quad (5)$$

where \mathcal{E}_r is the channel gain comprised of the gain due to the reflection/scattering and path-loss. $\tau_{rt} = 2d/c$ represents the round-trip delay of the reflected signal with c being the speed of light. The IF signal is sampled by an analog-to-digital converter (ADC) at the sampling rate, f_s , producing M_s samples per chirp per RF chain. Given the M_c chirps per frame and assuming a radar with M_a receiver antennas (with an RF chain for each antenna), a radar measurement produces $M_a M_c M_s$ ADC samples. We denote the ADC samples of each radar measurement (of a frame) by $\mathbf{X}^r \in \mathbb{C}^{M_a \times M_c \times M_s}$.

Further, a classical object detection method (e.g., [10]) is applied to the radar measurements, \mathbf{X}^r , to detect the objects as follows. (i) The range FFT, clutter cleaning, angle FFT, and Doppler FFT are applied to obtain a radar cube. (ii) The CFAR method detects the points with high power. (iii) These detected points are clustered with DBSCAN to determine the groups of detected points representing the candidate objects. (iv) The range o_k^r , angle o_k^a , and Doppler velocity o_k^v of each cluster k are estimated as the average of the cluster's points. Thus, for each object k , we have the properties $\mathbf{o}_k = [o_k^r, o_k^a, o_k^v]$. After this operation, we have K objects, which include the target.

III. PROBLEM DEFINITION: USER IDENTIFICATION WITH BEAM INFORMATION

In this work, we aim to determine the communication target within the radar measurement, which is an imperative task for leveraging radar sensing in communication systems. An important observation here is that this task requires a piece of communication information for the mapping between the radar and communication domains. For this purpose, we use the optimal beam index of the communication target, which may be obtained by initial access/channel estimation (mainly at sub-6 GHz) or beam sweeping (at higher frequencies). Then, given the beam index, we aim to find the communication user within the radar targets.

To formalize the problem, we first denote the number of samples by T and introduce the sub-index t to our notation. We then assume that there are K_t candidate objects in the radar measurement \mathbf{X}_t^r , which includes the communication target. Let us denote the index of this communication target by $k_t^c \in \{1, \dots, K_t\}$. Recall that for each sample $t \in \{1, \dots, T\}$, we have the radar measurements of the objects, $\mathbf{O}_t = [\mathbf{o}_{t,1}, \dots, \mathbf{o}_{t,K_t}]$, and the optimal beam index of the communication target, b_t^* . Given this information, our purpose is to determine the index of the communication target, k_t^c .

Mathematically, let us assume that there exists a function of parameters Θ that maps the radar measurement \mathbf{O}_t and the

optimal communication beam index b_t^* to the index k_t^c of the communication target among K_t radar-detected objects, i.e.,

$$f_{\Theta}(\mathbf{O}_t, b_t^*) = k_t^c. \quad (6)$$

Then, our aim becomes to approximate this function and parameters with minimal error, which can be formulated as

$$\hat{f}_{\Theta} = \arg \min_{f, \Theta} \frac{1}{T} \sum_{t=1}^T \mathcal{L}(f(\mathbf{O}_t, b_t^*), k_t^c), \quad (7)$$

for a given loss function \mathcal{L} . With this objective in mind, different approaches could be developed. For instance, a machine learning model can be trained using (7) directly. As an alternative, classical signal processing-based solutions may be proposed via the design of a fixed f_{Θ} function. In the following section, we present our solutions.

IV. PROPOSED SOLUTIONS

For the user identification problem, we develop two approaches: (i) A model-based baseline solution that maps the angle of the object in the radar to the angle of the communication beam, and (ii) a scalable machine learning solution that is agnostic to the number of candidate objects.

Baseline Solution: A key observation on the communication beamforming with the directional beams (e.g., DFT codebook) is that they point to the quantized angle of the communication object. Considering a line-of-sight channel, this information can be directly mapped to the radar angle. In this mapping, however, the angular misalignment of the radar and communication antennas may cause a significant error. Thus, we design a radar-to-communication angle mapping solution. In this solution, the angle corresponding to the communication beam index is mapped to the radar angle of the communication target, and the candidate object with the closest angle is selected as the communication target.

To formalize, we denote $\psi_{t,k}^r = o_{t,k}^a$ and $\psi_{t,k}^c$ as the radar and communication angles of the k -th candidate object in sample t . We also denote the pointing angles of the communication beams by $\phi_b, \forall b \in \{1, \dots, B\}$. With the proposed model and notation, the communication angle of the target object $\psi_{t,k_t^c}^c$ will have the minimum difference from the communication beam angle, as selecting this beam maximizes the received SNR. Therefore, we take $\psi_{t,k_t^c}^c \approx \phi_{b_t^*}$. This approximation particularly holds with a large number of narrow beams.

To convert the communication angles of the candidate objects to the radar angles, we adopt a simple model based on the angle offset (misalignment) between the radar and communication antennas. If ψ_0 denotes this angle offset, the transformation between the communication and radar angles of an object can be written as

$$\psi_{t,k}^r = \psi_0 + \psi_{t,k}^c, \quad \forall k \in \{1, \dots, K_t\}. \quad (8)$$

To estimate the offset, ψ_0 , we utilize mean squared error (MSE) over the data samples of the communication targets. Finally, to determine the index of the communication target, we select the candidate object that has the smallest distance to the radar angle of the target object, i.e.,

$$\hat{k}_t^c = \arg \min_{k \in \{1, \dots, K_t\}} |(\psi_0 + \phi_{b_t^*}) - \psi_{t,k}^r|. \quad (9)$$

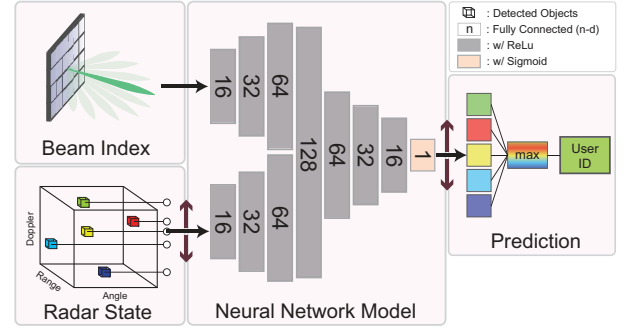


Fig. 2. The proposed DNN solution. Along with the communication beam index, the radar estimated state of each user is respectively fed to the DNN. The resulting outputs are collected together to select the communication target.

We note that the approach developed in this section assumes a line-of-sight channel and considers no other error than the misalignment of the radar and communication antennas. This approach could be expected to perform well in simulation-based models. In real systems, however, various imperfections (e.g., beam gain response of the actual hardware), in addition to the potential angle detection errors of the radar targets, may cause the approach to perform poorly. In the following, we aim to develop a robust solution that can adapt to these errors.

Deep Learning Solution: As mentioned, different types of imperfections, including object detection errors, may be present in real-world systems. To develop a robust solution, in this section, we propose a deep neural network model. This way, the imperfections can be learned and accommodated by the DNN models. Further, the additional information that has not been utilized in the baseline, i.e., the range and Doppler velocity of the objects, can be beneficial. Although the DNNs adapt to the imperfections by taking advantage of the available data, their design for this problem is not a straightforward task.

There are two important challenges with the design: (i) The number of candidate objects, K_t , is a sample-dependent parameter and can have various values at different instances. For this, one can design a model taking a predetermined maximum number of given objects as the input and append zeros to have an input of the predetermined size. Such an approach, however, could be a burden on the model's complexity. (ii) A desirable property of such a solution is permutation invariance. A solution that takes the information of all objects at once as an input may also utilize the ordering information of the objects, which is not desirable. Although this may be resolved by generating more samples by mixing the order of the objects, it possibly requires significantly more training for a large number of objects. *Overcoming both challenges, we propose our solution, where the radar information of each candidate object and the optimal communication beam index are respectively fed to the same neural network.* This network then returns the likelihood of the candidate object being the communication target. These soft predictions (likelihood ratios) are collected together for each candidate object, and the object with the maximum probability is returned as the target.

In this solution, the neural network takes four values as the input, i.e., the optimal beam index b_t^* , and range, angle, and Doppler velocity, $\mathbf{o}_{k,t}$. As shown in Fig. 2, in the architecture, we adopt a set of 7 layers with the given number of hidden

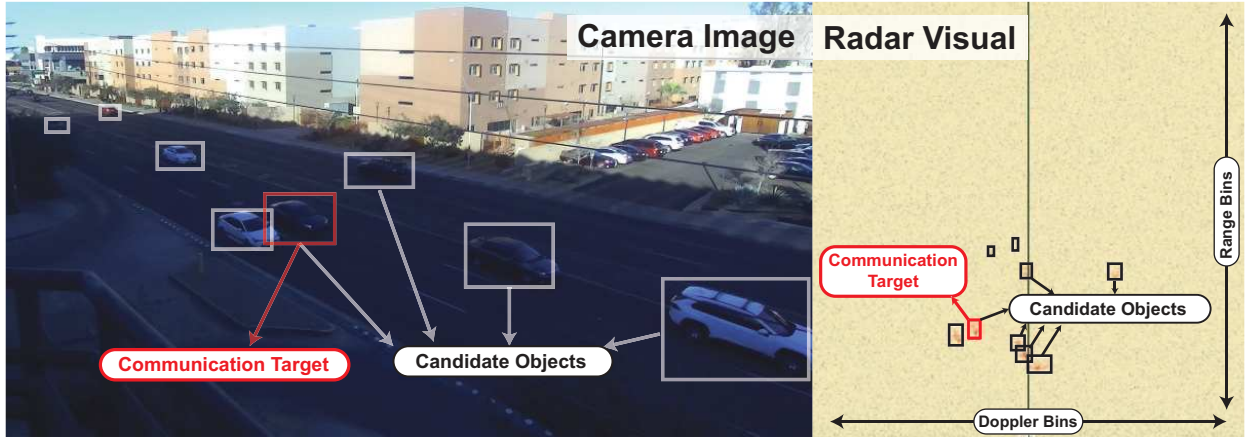


Fig. 3. A sample from the dataset is shown with the corresponding camera image and range-Doppler map. In this sample, the beam with the index 28 corresponds to the optimal beam. The targets are determined using this visual, based on a sequence of samples by tracking through time.

nodes, resulting in a total of 32,801 parameters. The network uses two parallel streams of three fully connected layers to extract features from radar data and beam indices. These features are then combined in a fourth layer. Subsequent layers reduce dimensionality for prediction. For the training of the model, we use $\arg \min_{\Theta} \frac{1}{T} \sum_{t=1}^T \sum_{k=1}^{K_t} \mathcal{L}(g_{\Theta}(\mathbf{o}_{t,k}, b_t^*), y_{t,k})$ where g_{Θ} represents the neural network function and constructs the function defined in (6) by $f_{\Theta}(\mathbf{O}_t, b_t^*) = \arg \max_k g_{\Theta}(\mathbf{o}_{t,k}, b_t^*)$. The variable $y_{t,k}$ is the indicator of the candidate object being the communication target, i.e., $y_{t,k} = 1$ if $k = k_t^c$ and 0 otherwise. From the objective, it can be seen that different candidate objects are treated as different samples. As the loss function, we utilize the MSE loss function.

For the complexity of the model, we point out that it linearly scales with the number of candidate objects, as each object is processed as a separate input. Similarly, the proposed model can be scaled to multiple users, where candidate objects and each user's beam information can be utilized as different inputs, with the cost of linearly scaling the complexity with the number of candidate objects. These issues, however, can be partially mitigated through batch processing techniques.

V. DATASET

For a real-world evaluation of the proposed user identification solutions, we built a dataset using a hardware testbed with co-existing radar and wireless mmWave equipment, following the DeepSense dataset structure [9]. Then, by processing the collected measurements/raw dataset, we built our development dataset for user identification in radar signatures. In this section, we describe our testbed and development dataset.

Testbed: We adopt Testbed 5 of the DeepSense 6G dataset [9] for the data collection, similar to [3], [4]. Testbed 5 comprises two units: (i) Unit 1, a fixed receiver acting as a basestation, and (ii) Unit 2, a mobile transmitter representing the target object. The rest of the details can be found in [3]. Differently from this work, the radar chirp parameters are selected for long-range performance, providing a maximum range of 249m and velocity of 82 km/h. The chirp frame covers $B = 310$ MHz bandwidth with a slope of $\mu = 10$ MHz/ μ s over $L = 250$ chirps/frame and $S = 512$ samples/chirp.

Development Dataset: We construct Scenario 35 of the DeepSense dataset [9]. In this scenario, a base station is placed on the second floor of a parking structure, directed towards a road with dense traffic. With this placement, we aimed to create a challenging data collection scenario representative of base station deployments. During the collection, a car with the transmitter is driven through the road in both directions. The received power via each beamforming vector and radar measurements are saved continuously at the rate of 9 samples/s, to be processed later. For the construction of the dataset, the index of the beam providing the highest received power is saved as the optimal beam index (b_t^*).

During processing, we used the optimal beam's power to identify samples where the communication target was within the scene and in line-of-sight, ensuring radar detectability. Specifically, long sequences containing sufficient receive power and a roughly linear optimal beam index pattern are included in the dataset. This intermediate dataset without labels contained 3045 samples. For the labels of the target objects, we applied a classical object detection solution. We then visually inspected the radar maps along with the RGB images. Through the evaluation of the samples over time, the objects that could be certainly identified as the target objects are marked. The final dataset contained 2158 samples. To prevent overfitting to specific trajectories, the 80/20% training/test split was performed on unique passes of the communication target through the radar's field of view, rather than on individual samples. There are 43 unique passes, leading to 34/9 passes and 1581/381 samples in the training/test sets.

Generalization: We highlight that both the baseline and proposed DNN models are site- and equipment-specific, depending on factors like antenna placement and alignment. Therefore, accurately identifying the communication target's index is essential, particularly when deploying to new environments or with different equipment. While this work uses manual radar data annotation for labeling, practical applications require more scalable methods. One promising technique is to have the target transmit unique radar-frequency signals (e.g., downchirps versus upchirps), allowing the receiver to identify the target alongside standard radar reflections. We leave the development of such methods to future work.

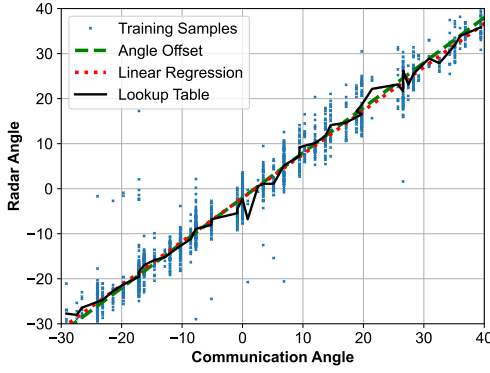


Fig. 4. The visualization of the data samples and the baseline solutions. Most of the samples follow a linear pattern with a relatively large variance.

VI. RESULTS

In this section, we evaluate the proposed solutions. In addition, we generalize the baseline solution and include three additional approaches: (i) Angle-based linear regression, where we use $\psi_{t,k}^r = \psi_0 + \alpha\psi_{t,k}^c$ for a parameter α instead of (8), (ii) linear regression for the whole radar state (range, angle, and Doppler velocity), and (iii) beam-radar angle lookup table, where each beam is mapped to the average radar angle of training samples, and the candidate object with the closest angle to the beam angle is selected as the target. For these, we estimate the parameters by utilizing the training set. For training the DNN solution, we employed the ADAM algorithm with a learning rate of 0.001. A learning rate step decay of 0.1 was applied at epochs 50 and 80. The network was trained for 100 epochs using a batch size of 128. To observe the impact of the small dataset size, we implemented 10-fold cross-validation. Within each fold, the training data was further partitioned into 80% training and 20% validation subsets. The model's average performance on these sets is also reported ¹.

First, in Fig. 4, we illustrate the radar and beam angles of the samples in our dataset, alongside the performance of baseline approaches. The data samples exhibit an average linear trend, supporting the design of linear models. However, we observe variations in the samples, likely due to detection inaccuracies, the radar's limited angular resolution, and other real-world errors. These variations can lead to performance degradation, particularly for linear solutions. Consequently, we anticipate that the DNN solution, with its ability to adapt to non-linear patterns, will achieve superior performance.

In Table I, we present the user identification accuracy among the candidates. The angle-based linear regression and angle offset solutions yield similar results, primarily highlighting the offset as the dominant challenge. Incorporating Doppler and range information offers a marginal improvement to linear regression. Similarly, the lookup table approach demonstrates comparable performance, suggesting that angular error is largely independent of the communication angle. The DNN solution achieves a 15 – 25% improvement across different sets, with a low cross-validation standard deviation, demonstrating

¹The implementation is available at <https://github.com/umut-demirhan/radar-user-identification>

TABLE I
USER IDENTIFICATION PERFORMANCE

Solution	Test Data	10-Fold Cross Validation	
	Accuracy	Avg. Acc.	Std. Dev.
Angle Offset (Baseline)	64.24	69.06	8.31
Linear Regression (Angle)	63.78	69.01	7.03
Linear Regression (3D)	67.43	70.82	8.34
Lookup Table	64.69	69.33	6.88
Deep Learning	89.29	85.81	4.14

the significance of capturing non-linear effects. Notably, the DNN performs slightly better on the test set compared to the cross-validation results, which is likely attributable to the smaller dataset size. Given that user identification is tightly coupled with potential radar-aided communication solutions, errors in this stage can be amplified throughout the system, potentially rendering the approach infeasible. Therefore, the high accuracy of our proposed deep learning approach is crucial for enabling practical radar-aided communication solutions.

VII. CONCLUSION

Radar-aided communications can be essential in advancing the performance of communication systems. One particular problem in aiding the communication with the radar is determining the relevant radar data to the communication target. In this paper, we formulated this problem and developed alternative solutions, including a DNN-based scalable approach. The DNN solution achieved 89% accuracy on the test set by significantly outperforming the baseline solutions and presented a promising result for enabling radar-aided communication.

REFERENCES

- [1] M. Alrabeiah, A. Hredzak, and A. Alkhateeb, "Millimeter wave base stations with cameras: Vision-aided beam and blockage prediction," in *VTC2020-Spring*. IEEE, 2020, pp. 1–5.
- [2] A. Ali, N. González-Prelcic, and A. Ghosh, "Millimeter wave V2I beam-training using base-station mounted radar," in *RadarConf*. IEEE, 2019, pp. 1–5.
- [3] U. Demirhan and A. Alkhateeb, "Radar aided 6G beam prediction: Deep learning algorithms and real-world demonstration," in *WCNC*. IEEE, 2022, pp. 2655–2660.
- [4] —, "Radar aided proactive blockage prediction in real-world millimeter wave systems," in *ICC*. IEEE, 2022, pp. 4547–4552.
- [5] F. Liu, W. Yuan, C. Masouros, and J. Yuan, "Radar-assisted predictive beamforming for vehicular links: Communication served by sensing," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7704–7719, 2020.
- [6] C. Aydogdu, F. Liu, C. Masouros, H. Wymeersch, and M. Rydström, "Distributed radar-aided vehicle-to-vehicle communication," in *RadarConf*. IEEE, 2020, pp. 1–6.
- [7] Z. Wang, K. Han, J. Jiang, F. Liu, and W. Yuan, "Multi-vehicle tracking and ID association based on integrated sensing and communication signaling," *IEEE Wireless Commun. Lett.*, vol. 11, no. 9, pp. 1960–1964, 2022.
- [8] U. Demirhan and A. Alkhateeb, "Integrated sensing and communication for 6G: Ten key machine learning roles," *IEEE Communications Magazine*, vol. 61, no. 5, pp. 113–119, 2023.
- [9] A. Alkhateeb, G. Charan, T. Osman, A. Hredzak, J. Morais, U. Demirhan, and N. Srinivas, "DeepSense 6G: A large-scale real-world multi-modal sensing and communication dataset," *IEEE Communications Magazine*, vol. 61, no. 9, pp. 122–128, 2023.
- [10] X. Li, X. Wang, Q. Yang, and S. Fu, "Signal processing for TDM MIMO FMCW millimeter-wave radar sensors," *IEEE Access*, vol. 9, pp. 167 959–167 971, 2021.