DeepSense 6G: A Large-Scale Real-World Multi-Modal Sensing and Communication Dataset

Ahmed Alkhateeb, Gouranga Charan, Tawfik Osman, Andrew Hredzak, João Morais, Umut Demirhan, and Nikhil Srinivas

Abstract—This article presents the DeepSense 6G dataset, which is a large-scale dataset based on real-world measurements of co-existing multi-modal sensing and communication data. The DeepSense 6G dataset is built to advance deep learning research in a wide range of applications in the intersection of multi-modal sensing, communication, and positioning. This article provides a detailed overview of the DeepSense dataset structure, adopted testbeds, data collection and processing methodology, deployment scenarios, and example applications, with the objective of facilitating the adoption and reproducibility of multi-modal sensing and communication datasets.

I. INTRODUCTION

The synergy between communication, multi-modal sensing, and positioning is envisioned as a defining characteristic of future wireless systems in 6G and beyond [1]–[8]. These systems will likely either implement co-existing communication and sensing functions or utilize one to aid the other. This is particularly motivated by the move to higher frequency bands where the wide bandwidth and large antenna arrays are attractive features for both communication and sensing. This synergy has recently been the driver for key research directions such as multi-modal sensing-aided communication [4]–[9], integrated sensing and communication [1], [3], [7], and communication-aided positioning [10]. Machine learning and deep learning have the potential to play fundamental roles in many of these problems [4]–[6], [11]. The ability, however, to develop and evaluate deep learning solutions depends on the availability of large enough datasets.

Multiple synthetic datasets have recently been developed with the objective of facilitating machine learning research in wireless communications. For example, in [12], the authors developed the DeepMIMO dataset, which is a parametric data generation framework based on ray-tracing simulations targeted mainly for MIMO research. In [13], the ViWi dataset was developed to generate both wireless and visual/LiDAR data. While providing a path for initial algorithm development, synthetic datasets may not be sufficient for obtaining realistic insights about the performance of the machine learning algorithms in real-world deployments.

With this motivation, we developed the DeepSense 6G dataset 1, the world’s first large-scale real-world multi-modal sensing and communication dataset. The DeepSense 6G dataset is (i) a large-scale dataset of more than 1 million data points, (ii) based on real-world measurements. The dataset comprises co-existing and synchronized multi-modal sensing and communication data and is organized in a collection of 40+ scenarios that cover diverse deployment use cases such as vehicle to infrastructure (V2I), vehicle to vehicle (V2V), pedestrians, drone communication, fixed-wireless, and indoor use cases. With that, the DeepSense 6G dataset has the potential to enable a wide range of applications in the intersection of communication, sensing, and positioning.

The goal of this article is to provide an overview of the DeepSense dataset structure, testbeds, data collection methodology, scenarios, and enabled applications. In particular, the article describes DeepSense in detail to facilitate the adoption and reproducibility of the DeepSense testbeds, data collection process, and datasets.

II. WHY DEEPSENSE 6G?

As discussed in the introduction, machine and deep learning are rapidly finding applications in the communication, sensing, and positioning systems and their intersection. The ability, however, to develop and adequately evaluate the performance of machine/deep learning approaches and derive meaningful conclusions is conditioned on having a dataset with the following key characteristics:

- **Co-existing sensing and communication**: To enable the targeted research directions, this dataset should comprise co-existing sensing and communication data that are synchronized to a sufficient extent. For example, both the wireless communication and sensing measurements need to be collected within the same coherence time.

- **Multi-modal sensing data**: By sensing, we do not mean only radar, but more broadly multi-modal sensing. This includes, for example, visual data, LiDAR sensory data, GPS positions, and weather data captured by sensors deployed at the infrastructure or mobile devices. Recent research has shown that a very interesting synergy exists between multi-modal sensing and communication [4]–[6], [11], which can be leveraged to benefit both of them.

- **Real-world measurements**: While using synthetic data, e.g. [12], [13], can provide initial insights, advancing the machine learning research and development towards real-world adoption and deployment requires the availability of real-world datasets. This is essential as some real-world and practical imperfections are very hard to

---

1 The DeepSense 6G dataset is publicly available at https://deepsense6g.net. Some of the DeepSense 6G scenarios (datasets) have been used before to enable the deep learning research in [6], [11], [14].
model and capture in the synthetic datasets. For example, practical positioning sensors do not normally provide accurate enough positions for narrow beam alignment and real-world visual/camera data may be sensitive to lighting/weather conditions.

- **Large-scale**: Developing deep learning solutions that are scalable and robust to data distribution shifts (due to changes in the environment or deployment) requires the availability of a large-scale dataset. Therefore, the ideal dataset should have a large number of data points and should have sufficient variance in both the communication and the multi-modal sensing measurements.

- **Scalable to various scenarios**: The applications and use cases of integrated multi-modal sensing and communication are countless. From the deployment use cases (V2I, V2V, indoor or outdoor, etc.) to the adopted devices (sub-6GHz massive MIMO, mmWave, reconfigurable intelligent surfaces (RIS), etc.), it is not feasible to have one single scenario that encompasses all the different applications. Therefore, the dataset definition and structure have to be scalable to enable the continued growth of the dataset.

To achieve these objectives, we built the DeepSense 6G dataset, which is a large-scale real-world dataset comprising co-existing multi-modal sensing and communication data covering various use cases and enabling many applications in the interplay of communications, sensing, and positioning.

### III. DeepSense Dataset Structure

The DeepSense 6G dataset is a *collection of scenarios*. Each scenario is a standalone dataset comprising multi-modal sensing and communication data collected in one, typically long, data collection session and formatted in a generic way that is unified across all the scenarios. The data collection of each scenario is planned to cover an important deployment scenario and enable one or more applications (the available scenarios will be discussed in Section VI). This simple structure achieves some important objectives: (i) It enables the scalability of the DeepSense datasets as it can keep growing to address more use cases by adding more scenarios. (ii) By providing data in a unified format for each scenario, it becomes easier to merge data from multiple scenarios, such as from different locations or at various times of the day. This approach enables the creation of a larger and more diverse dataset that meets specific objectives for targeted applications. (iii) It simplifies the access, the design, and the reproducibility of the dataset, which are critical goals for a useful deep learning dataset.

#### Data structure in each scenario:

Each DeepSense 6G scenario has the multi-modal sensing and communication data collected in one data collection session. As will be described in Section IV, each scenario uses a testbed that consists of a number of units (for example, a base station and a mobile user unit in a V2I scenario). Each unit is equipped with a number of sensors such as wireless receivers, radar, LiDAR, cameras, and GPS receivers. The data collection and processing are designed such that the data from all the sensors and units are collected and synchronized to the same sampling interval. Therefore, as shown in Fig. 1, the scenario data is eventually structured as a sequence of data groups, where each data group has the sensory data collected from all the sensors at the same sampling interval. To facilitate accessing the dataset, the DeepSense 6G website dedicates a page for each scenario with a detailed description of the adopted testbed, deployment, available modalities, and scenario folder structure. Additionally, the website has tutorials and videos to help researchers use the dataset, explaining everything from the dataset format to specific processing routines and visualizations.

### IV. DeepSense Testbeds

In order to ensure that the dataset is scalable and can cater to a diverse range of applications, we adopt a series of testbeds. These testbeds are developed to target specific deployment use cases such as V2I, V2V, and drone communication. As such, each testbed is designed to enable a particular application and is used to collect data for one or more scenarios that are relevant to the same application. For example, Testbed 1 was used to collect data for multiple DeepSense 6G scenarios (1 – 15), all related to V2I deployment. Each DeepSense scenario consists of data collected from a single data collection session.
session using only one of the testbeds and includes the multi-modal collected from the different sensors in the testbed. We adopt a unified and modular approach in building and describing these testbeds. Each testbed is composed of a number of units; each unit is a self-sustained module equipped with: (i) A set of wireless and environmental sensors such as wireless communication transceiver in one or more bands (such as mmWave phased arrays or sub-6GHz MIMO). The units are also equipped with a suite of sensors, such as an RGB camera, 2D/3D LiDAR, radar, and GPS RTK kit, to capture additional information about the surrounding environment. (ii) A processing unit to initialize and configure the sensors as well as store the collected sensing data. Further, the choice regarding the number of units in each testbed and the particular sensors to be deployed in each unit is based on the targeted application(s). Table I provides a summary of the testbeds and their adopted sensors.

Example - DeepSense Testbed 5: Given the modular design of the testbeds, providing one example may help clarify the testbed capabilities and enable the reproduction of the testbeds and datasets. For this objective, we now provide more details about a representative testbed, namely Testbed 5 developed for V2I applications. We designed this testbed to have two units: Unit 1 (a stationary unit), which acts as the basestation, and unit 2 (a vehicle), which represents the mobile user. Unit 1 is equipped with the following devices:

- **A mmWave receiver** that includes a 60GHz RF frontend with a 16-element uniform linear (phased) array (from SIVERS semiconductors). It adopts a combining beamforming codebook of 64 beams that uniformly scan 90° field-of-view.
- **RGB Camera** with 110° field-of-view, 30 frames-per-second (ZED2 from StereoLabs), and can sufficiently resolve objects of reasonable size (e.g., people and cars) at a distance of 100-150 m.
- **3D LiDAR** with 32 vertical and 1024 horizontal channels and with 120 m range and 20 Hz frame rate (from Ouster).
- **Radar** that adopts frequency modulated continuous wave (FMCW). It operates at 76-81GHz and has a maximum range of 100 m (from Texas Instruments).
- **GPS receiver** that uses real-time kinematic positioning (RTK) technology and a 10 Hz rate (from SparkFun).

The mobile user (unit 2) employs a mmWave transmitter with a 60GHz quasi-omni antenna (from SIVERS semiconductors) and the same GPS device used by the basestation.

In our experiments, we have evaluated the transmission range of the adopted mmWave transmitter/receiver and found that a 400 m can be reliably supported. This example testbed 5 has been used to collect scenarios 31 – 34. Each scenario dataset includes two sub-folders (named unit 1 and unit 2) containing the data captured by each unit. For example, the unit 1 sub-folder consists of the RGB image, radar I/Q samples, LiDAR point cloud, and the 64-element power vectors. Meanwhile, the unit 2 sub-folder contains the user’s GPS location.

V. DATA COLLECTION AND PROCESSING FRAMEWORK

As described in Section III, the DeepSense 6G dataset is a collection of scenarios, where each scenario consists of the multi-modal sensing and communication data collected in one long data collection session. Building the dataset of a new DeepSense 6G scenario comprises two main phases, i.e., data collection and data processing. Next, we present an overview of the two phases. Fig. 2 illustrates the different steps involved in generating the final development dataset.

A. Data Collection

The data collection process comprises two main phases: Off-field planning and on-field collection. The off-field planning phase primarily involves clearly defining the set of objectives, such as the number of units needed and the different sensors that must be deployed in each unit, and deciding on the location. The on-field data collection phase involves deploying the DeepSense 6G testbed in the pre-determined location.

**FoV alignment:** The different sensors, such as the mmWave phased array, RGB camera, LiDAR, and radar, have varying field-of-view (FoV) and range. For example: in the DeepSense testbed 5, the RGB camera has an FoV 110°, whereas the phased array has 90° FoV. In order to achieve a high-quality multi-modal dataset with multiple sensors, it is essential to align these different sensors before initiating the data collection process. For this, the testbed has provisions for some
Fig. 3. The DeepSense 6G dataset comprises 40+ scenarios with realistic datasets to encourage the development of novel applications. This figure presents the different deployments included in the DeepSense 6G dataset.

calibration steps to align the sensors considering the different ranges and FoVs.

**Capturing rate alignment:** The different sensing elements, such as the mmWave receiver, RGB camera, GPS RTK kit, LiDAR, and radar, can independently capture data at varying rates. For instance, the mmWave receiver runs at 10Hz, the camera can capture data at 30Hz, whereas the 3D LiDAR can complete a maximum of 20 sweeps in a second. Further, as described in Section IV, each of the DeepSense 6G testbeds is composed of multiple units, each equipped with an array of sensors capable of capturing data asynchronously. Therefore, data alignment across different modalities and multiple units is necessary to achieve meaningful data. We undertake a series of actions in order to achieve the necessary alignment. First, the software framework of the testbeds is designed to guarantee a comparable data collection rate for all modalities. For instance, the Testbed 5 presented in Section IV is designed and tested thoroughly to ensure the different modalities capture data at ≈10Hz, even though independently they can collect data at much higher rates. Additionally, the UTC timestamp of the data capture is stored for all modalities present in both units 1 and 2. Finally, during the post-processing, we align the different modalities to the time of mmWave receive beam sweeping by selecting the closest samples (to this time), resulting in coherent data for each time step.

**Real-time monitoring:** The final step in the on-field data collection is the real-time monitoring of the collected data. Apart from the manual check, we have also implemented data plotting capabilities in the testbed to enable us to monitor the data in real-time to ensure their quality and correctness.

**B. Data Processing**

The initial data collected using the testbed is referred to as the raw data. The raw data further undergoes a post-processing pipeline to generate the final development dataset. The post-processing steps include data synchronization and filtering. Next, we present the details of each of the steps.

**Data Synchronization:** The adopted sensing elements have different data capture rates, which might lead to an incoherent dataset. As described in Section V-A, the data capture flow in the testbed is designed to limit such misalignment. However, the provision in the testbed alone might not be sufficient. To avoid such inconsistency in the final dataset, the testbed is designed to record the UTC timestamp along with the actual captured data. The saved timestamps act as the pivot point and help in aligning the different data modalities. In particular, we consider the time instant of mmWave receive beam sweeping as the reference point and align the other modalities accordingly by selecting the samples that are closest in time. For this, we have developed scripts that can utilize the saved timestamps and generate a synchronized and aligned multi-modal dataset.

**Data Filtering:** The DeepSense 6G scenario datasets are collected to target specific deployment use cases and applications. Hence, an optional data filtering step may be necessary depending on the particular scenario (and the targeted application). It is important to note that during the data collection process, the testbed will continuously capture data until it is manually stopped. As a result, in scenarios involving a mobile user, some data samples may have the user outside the basestation’s FoV. For instance, in the V2I scenarios (such as scenarios 1–9) focused on sensing-aided beam prediction, only data samples with the mobile user within the basestation’s FoV (unit 2) are preferred. In such cases, a filtering process is


| Testbed | Scenarios | Deployment | Data Modalities | No. of Samples | Example Applications |
|---------|-----------|------------|-----------------|----------------|----------------------|
| 1       | 1-7       | V2I        | RGB images, GPS, mmWave receive power | 14,322         | Beam prediction [14], user identification, positioning |
|         | 8         |            | RGB images, GPS, mmWave receive power, and 2D LiDAR measurements | 4,043          | Beam prediction [8] |
|         | 9         |            | RGB images, GPS, mmWave receive power, 2D LiDAR and radar measurements | 5,964          | Beam prediction [7] |
|         | 10-15     | Pedestrian | RGB images, GPS, mmWave receive power | 1,528          | Beam and blockage prediction, user identification, positioning |
| 2       | 16        | Fixed Wireless | RGB images, GPS, mmWave receive power | 9,827          | Blockage identification and prediction |
| 3       | 17-22     | Fixed Wireless | RGB images, GPS, mmWave receive power | 460,838        | Blockage identification and prediction [4] |
|         | 24-29     |            | RGB images, GPS, mmWave receive power, 2D LiDAR | 500,000 | Blockage identification and prediction, object detection |
|         | 30        |            | RGB images, GPS, mmWave receive power, 2D LiDAR, and radar measurements | 14,660 | Blockage identification and prediction [11], object detection |
| 4       | 23        | Drone Comm. | RGB images, GPS, mmWave receive power | 11,387         | Beam prediction |
| 5       | 31-34     | V2I        | RGB images, GPS, mmWave receive power, 3D LiDAR, and radar measurements | 18,667         | Beam prediction [6], user and blockage identification |
|         | 35        |            | RGB images, GPS, mmWave receive power, and radar measurements | 3,045          | Beam prediction, user and blockage identification |
| 6       | 36-39     | V2V        | RGB images, GPS, mmWave receive power, 3D LiDAR, and radar measurements | 100,000        | Beam prediction, user and blockage identification |
| 7       | 40        | RIS        | RGB images, GPS, mmWave receive power, 3D LiDAR, and radar measurements | 5,000          | Beam prediction, user and blockage identification |

necessary to ensure that the final development dataset includes only the desired data samples. However, for scenarios designed primarily to study indoor/outdoor future blockage prediction (scenarios 16 – 22), which feature two stationary units, this optional filtering stage may not be necessary.

Data Verification: The final step of generating the development dataset is verifying the synchronized and filtered data. For that, we have developed in-house tools that assist in this verification process. One such tool is the data visualizer, which plots all the data modalities in the same GUI. It helps to filter out synchronization issues between different modalities and identify any missing or corrupted data. Overall, these tools help us ensure and maintain the highest quality of the datasets.

VI. DeepSense 6G Scenarios

The DeepSense 6G dataset consists of real-world multi-modal datasets from collected multiple locations. As described in Section III, the data is provided as different scenarios where each scenario is one long data collection. It consists of 40+ scenarios (and continuously growing) with realistic and challenging datasets to encourage the development of novel applications and enable varied deployments with sufficient variance. In this section, we highlight the inherent diversity of these scenarios and emphasize the value they add to the development of sensing-aided wireless communication applications.

Multiple sensors: Each sensing modality has its advantages and limitations. For example, radar data is primarily suitable for uncrowded scenarios. Utilizing multiple sensing modalities, i.e., data fusion, is a promising way to overcome such limitations, i.e., develop robust solutions by simultaneously leveraging multiple data modalities. However, developing such robust solutions requires the availability of large-scale multi-modal datasets. The DeepSense 6G dataset takes the first step towards enabling the development of such novel solutions. It consists of 40+ publicly available real-world scenarios with co-existing communication and multi-modal data collected with a set of sensors such as RGB camera, 2D/3D LiDAR, radar, GPS RTK kit, or a subset of them.

Multiple locations: Data collection location plays a critical role in determining the quality of wireless and sensing data. In particular, both stationary and dynamic objects in the environment impact the performance of the 5G and beyond wireless communication systems. Therefore, each location was carefully selected to capture realistic scenarios that helped in collecting meaningful and diverse data. The 40+ scenarios in the DeepSense 6G dataset were collected in two countries (USA and Spain) with more than 15+ indoor and outdoor locations. The outdoor locations have diverse characteristics; from quiet parking lots to busy downtown streets, the scenarios feature varying densities of pedestrians and vehicles, their mobility speeds, and environment geometry. Similarly,
for the indoor locations, we increase diversity by selecting locations with different shapes and sizes, such as conference rooms, corridors, etc. Further, to study the impact of weather conditions on the sensing modalities, we collect data in sunny, cloudy, windy, and rainy weather conditions and at various temperature and humidity levels.

**Different times of the day:** Similar to the locations, the time of data collection plays an important role in determining the overall quality of the collected data. For example, cameras are susceptible to lighting conditions resulting in noisy images in low-light/dark scenarios. Such degradation in image quality can significantly impact the downstream tasks. Developing robust sensing-aided solutions requires the availability of large-scale real-world diverse datasets. For that, we collected data at different times of the day to incorporate such diversity in the DeepSense 6G dataset. In particular, scenarios 2, 4, 14, 17 – 19, 33, and 34 are all collected during the night, with plans to add more scenarios shortly. The remaining scenarios were all collected during the day but at different times to increase variance. With data collected from early morning to late night, the scenarios of the DeepSense 6G dataset cover the entire spectrum.

**Different types of UEs:** Variance in object classes, object instances, and object speeds are crucial features of a dataset. Object classes and speeds are expected to significantly impact several downstream tasks, such as sensing-aided beam prediction, blockage prediction, and handoff. Developing solutions that can adapt and generalize across different deployment sites require dataset consisting of different real-world object classes (types). To that end, the different scenarios in the DeepSense 6G dataset are designed to capture this inherent variability of real-world locations. To provide UE variability in the dataset, the scenarios capture people on their skateboards, bike, or scooters, traveling along the sidewalk, and crossing the street. Also, certain scenarios were collected on some of the busiest downtown streets, which helped capture various types of vehicles, such as sedans, SUVs, trucks, buses, etc., traveling at different speeds. All that diversity in traffic makes the dataset diverse from both sensing and wireless perspectives.

**Deployment scenarios:** The DeepSense 6G dataset covers a wide range of deployment scenarios. In Fig 3, we show some of the adopted deployment scenarios and the associated testbeds. In particular, as shown in this figure, the DeepSense 6G dataset includes the following deployment scenarios: V2I, drone communication, V2V, RIS, indoor use cases, and pedestrians, to name a few. The sheer range of the deployment use cases that the DeepSense 6G dataset incorporates highlights its potential for enabling various critical applications.

VII. **Enabled Applications and ML Tasks**

The DeepSense 6G dataset, with its diverse scenarios, enables a wide range of applications in the interplay of communications, sensing, and positioning. This includes, as examples, multi-modal sensing aided beam and channel prediction, proactive blockage prediction and handoff, waveform design for ISAC systems, radar/LiDAR object detection and classification, and mmWave-based positioning. To accelerate the machine learning research in these directions and to facilitate results reproducibility, benchmarking, and comparisons, we also provide, as part of the DeepSense 6G dataset, task-specific development datasets for a number of key applications. It is important to note here that thanks to the unified approach in collecting and structuring the data of the different scenarios, the task-specific datasets can be constructed via combining data from multiple scenarios. This enables advanced machine learning research in transfer learning, generalizability, scalability, robustness, and distribution shift analysis, among others.

**Example - Generalizability evaluation for ML-based beam prediction:** Now, we describe how to use the dataset in studying one important research problem, namely multimodal vision and position-aided beam prediction. In general, the objective of this beam prediction task is to develop a machine learning-based solution that can utilize the input RGB image of the wireless environment (captured at the basestation) and the location information of the user to predict the optimal beam indices with high fidelity. However, different from the simple scenario-based beam prediction solution, we aim to study a more challenging problem of generalization across different unseen scenarios: Can a machine learning model be trained on a scenario dataset and then successfully generalize to other scenarios the model has not seen before? Therefore the objective is to develop a machine learning-based solution that can adapt to unseen scenarios that are not part of the training dataset. We adopt the machine-learning model proposed in [14] with a slight modification. Instead of utilizing a convolutional neural network (CNN)-based model for extracting meaningful features from the visual data, we utilize an object detection model, YOLOv3 [15], to detect and extract the bounding-box center coordinates of the different objects of interest in the visual data. The bounding-box coordinates and the normalized position data are then provided as input to a 2-layered feed-forward neural network to predict the optimal beam indices from a pre-defined beamforming codebook.

**Performance:** To evaluate the performance of the proposed solution, we adopt scenarios 1 and 5 of the DeepSense 6G dataset that consists of diverse wireless, visual, and position data. They were collected at two locations and at different times of the day and in different lighting conditions, making them good candidates for studying this generalizability problem. The final development dataset consists of 2411 and 2300
samples from scenarios 1 and 5. It comprises pairs of visual and position data along with the ground-truth beam indices. The data from each scenario is further divided into training and validation sets following a split of 70% – 30%. Fig. 4 plots the confusion matrices for the following two cases: (a) Training and testing on scenario 5 development dataset and (b) training on scenario 1 data and evaluating on scenario 5 test dataset. This figure shows that even with no training data from scenario 5, the proposed machine learning-based solution can still predict the optimal beams with sufficient accuracy (more than 70% top-3 accuracy). Fig. 4 also shows that even when the model predicts a wrong beam index, this index is with high probability close to the optimal ground-truth beams and hence may still have good receive power. These results highlight the value of using the DeepSense 6G dataset to obtain practical insights about the performance of a developed machine learning solution in realistic communication environments.

VIII. CONCLUSION

This article presented the DeepSense 6G dataset, which is a large-scale real-world dataset comprising co-existing and synchronized multi-modal sensing and communication data from more than 40 deployment scenarios, including vehicle-to-infrastructure, vehicle-to-vehicle, reconfigurable intelligent surfaces, pedestrians, and drone communication. The article described the motivation behind the dataset, its scalable structure, the modular testbed definition, the available scenarios, and some of the machine learning applications enabled by the dataset. The goal of the DeepSense 6G dataset project is to advance the deep learning research and development in a wide range of applications in the interplay of communication and sensing to enable the reproduction and benchmarking of their datasets and research results.

REFERENCES

[1] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng, L. Zheng, and A. Petropulu, “An overview of signal processing techniques for joint communication and radar sensing,” IEEE JSAC, vol. 15, no. 6, pp. 1295–1315, 2021.

[2] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, “Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond,” IEEE JSAC, vol. 40, no. 6, pp. 1728–1767, 2022.

[3] A. Liu, Z. Huang, M. Li, Y. Wan, W. Li, T. X. Han, C. Liu, R. Du, D. K. P. Tan, J. Lu, Y. Shen, F. Colone, and K. Chetty, “A survey on fundamental limits of integrated sensing and communication,” IEEE Communication Surveys and Tutorials, vol. 24, no. 2, pp. 994–1034, 2022.

[4] G. Charan and A. Alkhateeb, “Computer Vision Aided Blockage Prediction in Real-World Millimeter Wave Deployments,” 2022 IEEE Globecom Workshops, 2022, pp. 1711–1716.

[5] M. Zecchin, M. B. Mashhadi, M. Jankowski, D. Gündüz, M. Kountouris, and D. Gesbert, “LiDAR and position-aided mmWave beam selection with non-local CNNs and curriculum training,” IEEE Transactions on Vehicular Technology, vol. 71, no. 3, pp. 2975–2990, 2022.

[6] G. Charan, U. Demirhan, J. Moraes, A. Behboodi, H. Pezeshki, and A. Alkhateeb, “Multi-modal beam prediction challenge 2022: Towards generalization,” 2022. [Online]. Available: https://arxiv.org/abs/2209.07519

[7] U. Demirhan and A. Alkhateeb, “Radar aided 6G beam prediction: Deep learning algorithms and real-world demonstration,” in IEEE WCNC, 2022, pp. 2655–2660.

[8] S. Jiang, G. Charan, and A. Alkhateeb, “LiDAR aided future beam prediction in real-world millimeter wave V2I communications,” IEEE Wireless Communications Letters, pp. 1–1, 2022.

[9] A. Ali, N. Gonzalez-Prelcic, and A. Ghosh, “Passive radar at the roadside unit to configure millimeter wave vehicle-to-infrastructure links,” IEEE Transactions on Vehicular Technology, vol. 69, no. 12, pp. 14903–14917, 2020.

[10] H. Kim, K. Granström, L. Gao, G. Battistelli, S. Kim, and H. Wyneersch, “5G mmWave cooperative positioning and mapping using multi-model PHD filter and map fusion,” IEEE Transactions on Wireless Communications, vol. 19, no. 6, pp. 3782–3795, 2020.

[11] U. Demirhan and A. Alkhateeb, “Integrated sensing and communication for 6G: Ten key machine learning roles,” 2022. [Online]. Available: https://arxiv.org/abs/2208.02157

[12] A. Alkhateeb, “DeepMIMO: A generic deep learning dataset for millimeter wave and massive MIMO applications,” in Proc. of ITA, San Diego, CA, Feb 2019, pp. 1–8.

[13] M. Abrabehia, A. Hredzak, Z. Liu, and A. Alkhateeb, “ViWi: A deep learning dataset framework for vision-aided wireless communications,” in IEEE VTC, 2020, pp. 1–5.

[14] G. Charan, T. Osman, A. Hredzak, N. Thawdar, and A. Alkhateeb, “Vision-position multi-modal beam prediction using real millimeter wave datasets,” in IEEE WCNC, 2022, pp. 2727–2731.

[15] J. Redmon and A. Farhadi, “YOLOv3: An incremental improvement,” 2018. [Online]. Available: https://arxiv.org/abs/1804.02767

Ahmed Alkhateeb is an Assistant Professor in the School of Electrical, Computer and Energy Engineering at Arizona State University (ASU).

Gouranga Charan is currently pursuing the Ph.D. degree in Electrical Engineering at Arizona State University, USA.

Tawfik Osman is currently pursuing the Ph.D. degree in Electrical Engineering with Arizona State University, USA.

Andrew Hredzak received the B.S. degree in Electrical Engineering from Arizona State University, USA, in 2021.

João Moraes is currently pursuing the Ph.D. in Electrical Engineering with Arizona State University, USA.

Umut Demirhan is currently pursuing the Ph.D. in Electrical Engineering with Arizona State University, USA.

Nikhil Srinivas received the M.S degree in Electrical Engineering from Arizona State University, USA, in 2022.