Sensor systems

M3-CFR: A Domain-Adaptive Bistatic mmWave MIMO CFR Dataset for Microgesture Recognition

Khandaker Foysal Haque*, Francesca Meneghello**, and Francesco Restuccia***

Institute for Intelligent Networked Systems, Northeastern University, Boston, MA 02115 USA

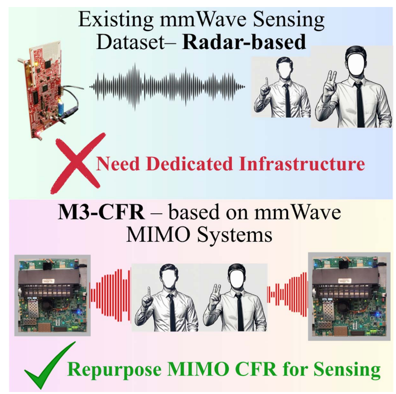
*Graduate Student Member, IEEE

**Member, IEEE

***Senior Member, IEEE

Manuscript received 21 February 2026; revised 1 May 2026; accepted 6 June 2026. Date of publication 15 June 2026; date of current version 13 July 2026.

Abstract—In this letter, we introduce M3-channel frequency response (CFR), a bistatic millimeter-wave (mmWave) multiple-input multiple-output (MIMO) CFR dataset for fine-grained microgesture recognition under domain shifts. As wireless systems move toward integrated sensing and communication (ISAC), sensing will increasingly be realized within communication infrastructures rather than via standalone sensing hardware. Yet most existing mmWave sensing datasets are radar-centric and sensing-specific, which misaligns with communication-driven ISAC deployment. In contrast, M3-CFR provides communication-native channel frequency response (CFR) from an 8×8 fully digital mmWave MIMO ISAC testbed at 58 GHz with 1 GHz bandwidth. It contains ten microgestures from three subjects across three indoor environments (3000 synchronized CFR instances). The dataset consists of raw and domain-adaptively processed CFR. The latter includes path-loss compensation, frequency normalization, and an entropy-guided subchannel filtering pipeline, enabling cross-environment and cross-subject sensing benchmarking.



Index Terms—Sensor systems, channel frequency response (CFR) gesture recognition, channel measurements dataset, multiple-input multiple-output (MIMO) domain adaptation, microgestures, mmWave sensing.

I. INTRODUCTION

Wireless sensing has become a core capability of next-generation communication systems, enabling diverse applications, including human activity recognition [1], [2], gesture recognition [1], and environmental awareness, without requiring dedicated sensing hardware. This trend is further reinforced by integrated sensing and communication (ISAC), where sensing functionalities are expected to be realized directly within communication infrastructures using communication-native signals. While significant progress has been made using sub-6 GHz Wi-Fi channel frequency response (CFR) and millimeter-wave (mmWave) radar systems, existing datasets largely rely on either sensing-specific hardware or low-frequency channels with limited spatial resolution.

mmWave multiple-input multiple-output (MIMO) communication systems offer a promising solution for fine-grained sensing due to their large bandwidth and high spatial diversity. In particular, the CFR, routinely estimated for MIMO operations, embeds rich spatial, temporal, and frequency-domain information that is sensitive to subtle changes in the physical world. Leveraging mmWave CFR aligns naturally with the ISAC paradigm, as it enables sensing without modifying transmission protocols or introducing additional sensing signals or devices. However, despite this potential, publicly available datasets providing high-resolution mmWave MIMO CFR measurements for human-centric sensing tasks remain limited. mmWave-based gesture recognition has been extensively investigated in prior work, including approaches based on radar-centric pipelines, dedicated sensing signals, and standard communication waveforms using nondedicated hardware. Representative examples include end-to-end radar-based frameworks leveraging Doppler and angle-domain representations [3], and multiview signal-based approaches for improved robustness [4]. While these works demonstrate strong performance, they are predominantly based on radar signals or dedicated sensing waveforms and focus on task-specific pipelines rather than releasing standardized, publicly available datasets. As a result, reproducible benchmarking of fine-grained gesture recognition using communication-native MIMO CFR for ISAC remains underexplored.

To address these limitations, we introduce M3-CFR, a bistatic mmWave MIMO CFR dataset designed for fine-grained microgesture recognition under domain shifts. M3-CFR data samples are collected using an 8×8 fully digital mmWave MIMO ISAC testbed [5] operating at 58 GHz with 1 GHz bandwidth, capturing communication-native MIMO CFR measurements during subtle hand gestures. The dataset comprises ten microgestures performed by three subjects across three indoor environments, yielding 3000 temporally synchronized CFR instances. In addition to raw measurements, we release a domain-adaptive CFR preprocessing pipeline and the corresponding processed data to support robust cross-environment and cross-subject learning. The main contributions of this work are summarized as follows.

Summary of Contribution:

1) We introduce a publicly available bistatic mmWave MIMO CFR dataset for fine-grained microgesture recognition within a communication-oriented ISAC framework, collected using an 8×8

Corresponding author: Khandaker Foysal Haque (e-mail: haque.k@northeastern.edu).

Associate Editor: Filippo Costa.

Digital Object Identifier 10.1109/LSENS.2026.3703644

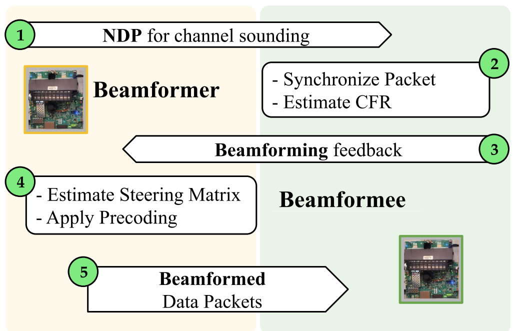


Fig. 1. CFR estimation through channel sounding in MIMO.

fully digital mmWave MIMO testbed [5] operating at 58 GHz with 1 GHz bandwidth.

2) The dataset comprises ten microgestures performed by three subjects across each of three indoor environments, yielding 3000 temporally synchronized high-resolution CFR instances, enabling systematic evaluation of cross-environment and cross-subject domain shifts.

3) We release both the raw and preprocessed CFR measurements and a domain-adaptive preprocessing pipeline incorporating path loss compensation, frequency normalization, and entropy-guided subchannel filtering to enhance robustness under realistic deployment variations.

4) We develop a temporal convolutional network (TCN) and convolutional neural network (CNN)-based microgesture recognition baselines to establish a consistent reference framework for future learning-based and domain-adaptive sensing research using the dataset. We share the code and dataset online.[1]

II. BACKGROUND AND RELATED WORK

A. Channel Estimation in MIMO

Modern MIMO wireless systems estimate the channel state through explicit channel sounding procedures in order to support spatial multiplexing and beamforming. Fig. 1 illustrates a generic bistatic MIMO sounding and beamforming workflow, which forms the basis for leveraging the CFR as a communication-native sensing primitive.

CFR estimation with channel sounding: As shown in *Step 1* of Fig. 1, the transmitter (TX) (referred to as the *beamformer*) periodically broadcasts a known training signal in null data packet for channel sounding. Upon reception (*see Step 2*), the receiver (RX) synchronizes the packets and estimates the frequency-domain MIMO channel response, referred to as CFR, across the occupied bandwidth. The resulting channel estimate can be expressed as $\mathbf{H} \in \mathbb{C}^{K \times M \times N}$, where K denotes the number of frequency subchannels, and M and N denote the numbers of transmit and receive antennas, respectively. The estimated CFR captures the intrinsic propagation characteristics of the environment, free from interstream or interuser interference.

Beamforming and data transmission: The estimated CFR is sent back (see *Step 3*) to the beamformer. The beamformer uses the CFR to compute the precoding matrix (see *Step 4*) that allows focusing energy along dominant spatial paths. This enables subsequent beamformed data transmission (see *Step 5*) that can be used to improve system's robustness and receive signal-to-noise ratio, or to enable multiplexing. While these beamforming operations are essential for data delivery, the sensing information captured in this work is obtained *prior* to precoding, i.e., we are collecting the estimated CFR during the sounding phase.

CFR as a sensing primitive: MIMO CFR encodes rich spatial, temporal, and frequency-domain information, which is highly sensitive to environmental dynamics and human motion. Changes in body posture, hand gestures, or movement patterns induce subtle but measurable variations in multipath propagation, which are reflected in the estimated CFR across antennas and subchannels. Importantly, CFR measurements are natively available in communication systems and do not require sensing-specific waveforms, protocol modifications, or additional hardware.

[1][Online]. Available: https://github.com/Restuccia-Group/M3-CFR

TABLE 1. Comparison of Representative Wireless Sensing Datasets

Dataset	Band	Modality	Envs	Subj.	Stream
Widar3 [6]	Sub-6 GHz	CFR	3	16	MIMO
Human AR [7]	Sub-6 GHz	CFR	3	1	MIMO
Meneghello et al. [8]	Sub-6 GHz	CSI	7	13	MIMO
mmBody [9]	mmWave	Radar	7	20	–
MiliPoint [10]	mmWave	Radar	–	–	–
DISC [11]	mmWave	CIR	2	7	SISO
M3-CFR (Ours)	mmWave	CFR	3	3	MIMO

B. Existing Datasets

Existing wireless sensing datasets have predominantly focused on either sub-6 GHz Wi-Fi CFR or mmWave radar-based measurements, as summarized in Table 1.

Sub-6 GHz CFR-based sensing datasets: Sub-6 GHz CFR datasets, such as Widar3 [6], Human AR [7], and the dataset by Meneghello et al. [8], leverage commodity Wi-Fi devices and support activity or gesture recognition across multiple environments and subjects. However, their limited bandwidth and spatial resolution constrain sensitivity to fine-grained motions, making microgesture recognition challenging.

mmWave radar-based sensing dataset: To achieve higher resolution, several mmWave radar-based datasets have been introduced, including mmBody [9] and MiliPoint [10]. These datasets provide rich spatial information using dedicated sensing hardware and radar-specific waveforms, but their sensing pipelines differ fundamentally from communication systems, limiting direct applicability to communication-centric ISAC deployments.

mmWave ISAC sensing dataset: More recently, mmWave ISAC sensing datasets have emerged. DISC [11] provides channel impulse response (CIR) measurements extracted from standard-compliant mmWave transmissions. However, it is limited in MIMO dimensionality and domain diversity.

In contrast, M3-CFR focuses on communication-native mmWave MIMO CFR as the sensing primitive. By providing full-rank 8×8 CFR measurements at 58 GHz with synchronized video ground truth across multiple environments, M3-CFR enables systematic evaluation of cross-environment and cross-subject microgesture recognition within an ISAC framework.

III. M3-CFR DATASET DESCRIPTION

The M3-CFR dataset includes measurements corresponding to ten distinct microgestures— *thumbs up, thumbs down, victory, okay, pointing, pinch, tap fingers, fingers crossed, fist*, and *half-open palm* as presented in Fig. 2. These gestures were selected to represent subtle, fine-grained hand movements, which we refer to as *microgestures*. In this work, microgestures denote low-amplitude, localized finger or hand motions that induce only minor perturbations in the wireless channel, in contrast to coarse hand or arm movements that produce

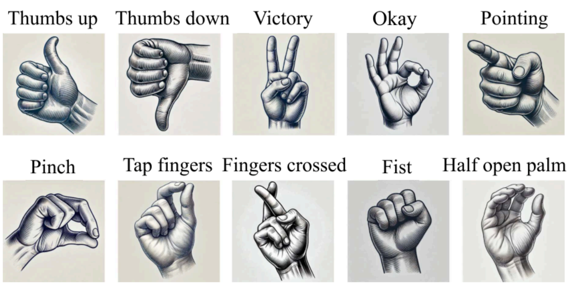


Fig. 2. Different microgestures considered in M3-CFR dataset: *thumbs up, thumbs down, victory, okay, pointing, pinch, tap two fingers, fingers crossed, fist, and half-open palm.*

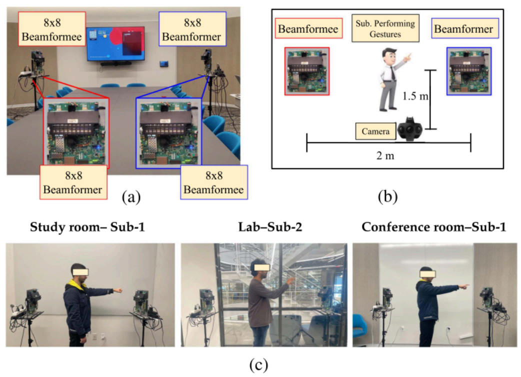


Fig. 3. Experimental setup and data collection overview. (a) m^{3}MIMO testbed [5]. (b) Experimental deployment. (c) Sample gesture frames across environments and subjects.

stronger and more easily detectable signal variations. Such microgestures are representative of realistic interactions and pose a more challenging sensing task, making them well-suited for evaluating the sensitivity and robustness of mmWave CFR-based sensing.

Data were collected from three human subjects across three indoor environments, namely, a laboratory space, a conference room, and a study room. Each subject performed each gesture 50 times per environment, with each gesture instance lasting 3 s. CFR samples were recorded at a rate of 166 packets/s, resulting in approximately 500 CFR frames per gesture instance.

Overall, the dataset contains a total of 3000 labeled gesture instances, corresponding to approximately 1.5 million CFR frames. Each gesture instance is represented as a 4-D tensor of size $T \times K \times N \times M$, where $T \approx 500$ denotes the number of temporal samples, $K = 1024$ is the number of orthogonal frequency-division multiplexing (OFDM) subchannels, and $N = 8$ and $M = 8$ correspond to the number of transmit and receive antennas, respectively. To enable accurate ground-truth annotation, high-resolution video recordings were captured synchronously with the CFR measurements using a fixed camera.

A. Experimental Setup

The M3-CFR dataset was collected using the m^{3}MIMO testbed [5], a fully digital 8×8 mmWave MIMO platform for ISAC as shown in Fig. 3(a). The testbed consists of two software defined radio nodes, each equipped with eight RF chains, enabling full-rank MIMO operation with digital beamforming. One node operated as the TX and the other as the RX, forming a bistatic configuration.

As illustrated in Fig. 3(b), the TX and RX were mounted at a height of $\sim$1.3 m and placed 2 m apart, with subjects positioned at the midpoint and performing single-hand microgestures under a fixed body posture to isolate gesture-induced channel variations. Singleuser-MIMO OFDM waveforms were transmitted at a center frequency of 58 GHz with 1 GHz bandwidth, enabling extraction of CFR measurements over 1024 subchannels. To provide ground-truth annotation, a high-resolution camera was placed laterally to the TX–RX axis and captured video synchronously at 30 fps. Representative video frames from different subjects and environments are shown in Fig. 3(c).

B. Domain-Adaptive CFR Preprocessing

CFR measurements in the mmWave spectrum are sensitive to propagation distance, multipath conditions, and spatial configuration, leading to domain-dependent distortions that hinder cross-environment and cross-subject generalization. These domain shifts manifest as amplitude scaling, spectral distortions, and unstable subchannel responses. To support domain adaptation, we apply a lightweight preprocessing pipeline comprising path loss compensation, frequency normalization, and entropy-guided subchannel filtering, which, respectively, reduce amplitude variations, suppress global power differences, and attenuate unstable subchannels, thereby better isolating gesture-induced variations that are more consistent across environments and subjects.

1) Path Loss Compensation: Large-scale path loss introduces amplitude scaling in CFR magnitudes due to variations in propagation distance and shadowing. We model the path loss as $P_L(d) \propto 10n\log_{10}(d) + \chi_\sigma$, where d is the TX–RX separation, n is the path loss exponent, and $\chi_\sigma \sim \mathcal{N}(0, \sigma^2)$ captures shadow fading. Each CFR sample is compensated using a distance-dependent factor $A_{\mathrm{PL}}(d) = d^{-n}$. For the CFR $\mathbf{H}_r^{m,n}(k)$ at time index r, the compensated CFR is $\mathbf{C}_{\mathrm{comp},r}^{m,n}(k) = \mathbf{H}_r^{m,n}(k) A_{\mathrm{PL}}(d)$.

2) Frequency Normalization: To suppress global power variations and emphasize relative spectral patterns, the compensated CFR is normalized using its Frobenius norm

$$\mathbf{C}_{\mathrm{norm},r}^{m,n}(k) = \frac{\mathbf{C}_{\mathrm{comp},r}^{m,n}(k)}{\|\mathbf{C}_{\mathrm{comp},r}\|} \tag{1}$$

where

$$\|\mathbf{C}_{\mathrm{comp},r}\| = \sqrt{\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{k=1}^{K}\left|\mathbf{C}_{\mathrm{comp},r}^{m,n}(k)\right|^2}. \tag{2}$$

This normalization reduces frame-to-frame and environment-dependent power discrepancies.

3) Entropy-Guided Subchannel Filtering: Subchannels affected by unstable multipath and noise degrade the reliability of gesture-relevant features. To emphasize stable subchannels, we compute the spatial entropy of each subchannel across MIMO links

$$E(k) = -\sum_{i=1}^{MN} p_i \log p_i, \quad p_i = \frac{\left|\mathbf{C}_{\mathrm{norm}}^{i}(k)\right|^2}{\sum_{j=1}^{MN}\left|\mathbf{C}_{\mathrm{norm}}^{j}(k)\right|^2} \tag{3}$$

where i indexes TX–RX antenna pairs. Lower entropy indicates more concentrated and stable energy distributions. Each subchannel is assigned a weight $w(k) = \exp(-E(k))$, yielding the filtered CFR, $\mathbf{C}_{\mathrm{filt},r}^{m,n}(k) = \mathbf{C}_{\mathrm{norm},r}^{m,n}(k)\, w(k)$.

IV. BENCHMARK TASK AND EVALUATION

We evaluate the M3-CFR dataset using microgesture recognition as a supervised multiclass classification task based on CFR measurements. Each gesture instance contains approximately 500 CFR measurements captured over a 3 s interval. For benchmarking, we segment each instance into shorter temporal windows of $T_w = 10$ consecutive

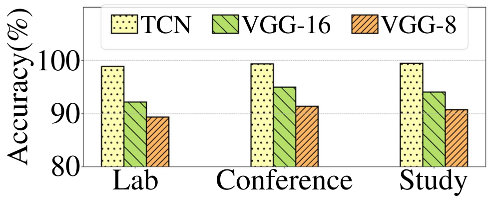


Fig. 4. Gesture recognition accuracy across different environments using TCN, VGG-16 and VGG-8.

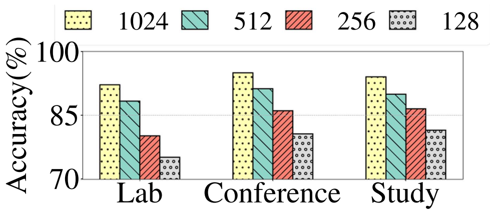


Fig. 5. Gesture recognition accuracy using VGG-16 under varying numbers of CFR subchannels across environments.

CFR frames, each assigned the corresponding gesture label. Each CFR snapshot is first reshaped from $K \times N \times M$ to $K \times S$, where $K = 1024$ denotes the subcarriers and $S = NM = 64$ corresponds to the spatial streams from the 8×8 antenna pairs. The complex-valued CFR is then represented by concatenating the real and imaginary components along the spatial dimension, resulting in a dimension of $K \times (2S) = 1024 \times 128$. Consequently, each input sample is represented as a tensor of size $T_w \times K \times (2S) = 10 \times 1024 \times 128$.

For dataset benchmarking, we adopt representative learning-based baselines based on both convolutional and temporal architectures. Specifically, we employ VGG-based [12] CNN models (VGG-8 and VGG-16) and a TCN [13] to capture complementary characteristics of CFR data. VGG-16 consists of 13 convolutional and three fully connected layers, while VGG-8 provides a reduced-depth variant with fewer layers and lower computational complexity. The TCN follows a lightweight architecture with stacked dilated causal convolutional layers (five layers with increasing dilation rates) and global pooling to model temporal dependencies in the CFR sequence. The objective is not to optimize model performance, but to provide consistent and reproducible baselines for evaluating the sensing quality of the dataset.

Fig. 4 compares gesture recognition accuracy across environments using TCN, VGG-16, and VGG-8. The TCN achieves near-perfect accuracy (98.9%–99.5%), outperforming VGG-16 (92.2%–95.0%) by 4%–7% and VGG-8 (89.3%–91.4%) by 8%–10%, highlighting the benefit of explicit temporal modeling. Among the convolutional baselines, VGG-16 consistently exceeds VGG-8 by $\sim$3%, indicating that deeper architectures better capture spatial structure, although they remain less effective than TCN in exploiting temporal dynamics.

Next, using the VGG-16 model, Fig. 5 evaluates gesture recognition performance as a function of the number of CFR subchannels. With the full 1024 subchannels, the model achieves accuracies of 92.19%, 94.99%, and 94.06% across the three environments. As the number of subchannels is reduced to 512, 256, and 128, accuracy degrades gradually, remaining above 88%, 86%, and 75%, respectively. This performance trend indicates that M3-CFR retains discriminative gesture information under bandwidth-limited conditions, supporting the robustness of communication-native mmWave CFR for microgesture recognition.

Limitations and future work: While M3-CFR captures fine-grained mmWave CFR measurements across multiple environments and subjects, the current release is limited to three indoor environments and three subjects due to the substantial time and logistical effort required for such large-scale data collection campaigns. Expanding the dataset to include additional environments, subjects, and deployment scenarios is an important direction for future work and will further strengthen cross-domain evaluation.

V. CONCLUSION

This work introduced M3-CFR, a bistatic mmWave MIMO CFR dataset for fine-grained microgesture recognition within an ISAC framework. By providing high-resolution 8×8 CFR measurements synchronized with visual ground truth across multiple environments and subjects, the dataset facilitates systematic investigation of cross-domain wireless sensing under realistic deployment conditions. In addition to raw measurements, we release domain-adaptively processed CFR data to facilitate mitigation of environment and subject induced variability. Baseline results using a CNN-based framework illustrate the feasibility of reliable microgesture recognition from communication-native mmWave CFR and highlight the role of subchannel resolution in sensing performance. Overall, M3-CFR establishes a communication-centric reference dataset for advancing learning-driven mmWave sensing within future ISAC systems. It provides a foundation for future investigations into domain generalization, robustness, and learning-driven wireless sensing in mmWave ISAC systems.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation under Grant ECCS-2229472, in part by the Air Force Office of Scientific Research under Grant FA9550-23-1-0261, and in part by the Office of Naval Research under Grant N00014-23-1-2221.

REFERENCES

[1] K. F. Haque, M. Zhang, F. Meneghello, and F. Restuccia, “BeamSense: Rethinking wireless sensing with MU-MIMO Wi-Fi beamforming feedback,” *Comput. Netw.*, vol. 258, 2025, Art. no. 111020.
[2] K. F. Haque, M. Zhang, F. Meneghello, and F. Restuccia, “Si-FI: Learning the beamforming feedback for simultaneous multi-subject sensing,” *Comput. Netw.*, vol. 275, 2025, Art. no. 111907.
[3] A. Ali et al., “End-to-end dynamic gesture recognition using mmwave radar,” *IEEE Access*, vol. 10, pp. 88692–88706, 2022.
[4] Q. Chen, Q. Lu, Y. Chen, Y. Tian, Z. Cui, and Z. Cao, “Domain-generalized gesture recognition via mmwave radar signal multiview learning,” *IEEE Trans. Instrum. Meas.*, vol. 74, pp. 1–14, 2025.
[5] K. F. Haque, F. Meneghello, K. M. Rumman, and F. Restuccia, “m3MIMO: An 8×8 mmWave multi-user MIMO testbed for wireless research,” in: *Proc. MobiCom*, 2024, pp. 1922–1929.
[6] Y. Zheng et al., “Zero-effort cross-domain gesture recognition with Wi-Fi,” in *Proc. ACM Int. Conf. Mobile Syst., Appl., Serv.*, 2019, pp. 313–325.
[7] A. Baha’A, M. M. Almazari, R. Alazrai, and M. I. Daoud, “A dataset for Wi-Fi-based human activity recognition in line-of-sight and non-line-of-sight indoor environments,” *Data Brief* vol. 33, 2020, Art. no. 106534.
[8] F. Meneghello, N. Dal Fabbro, D. Garlisi, I. Tinnirello, and M. Rossi, “A CSI dataset for wireless human sensing on 80 MHz Wi-Fi channels,” *IEEE Commun. Mag.*, vol. 61, no. 9, pp. 146–152, Sep. 2023.
[9] A. Chen, X. Wang, S. Zhu, Y. Li, J. Chen, and Q. Ye, “mmBody benchmark: 3D body reconstruction dataset and analysis for millimeter wave radar,” in *Proc. 30th ACM Int. Conf. Multimedia*, 2022, pp. 3501–3510.
[10] H. Cui et al., “MiliPoint: A point cloud dataset for mmWave radar,” in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 36, 2023, pp. 62713–62726.
[11] J. Pegoraro, J. O. Lacruz, M. Rossi, and J. Widmer, “DISC: A dataset for integrated sensing and communication in mmWave systems,” *IEEE Comm. Mag.*, vol. 63, no. 10, pp. 94–100, Jun. 2025.
[12] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014, *arXiv:1409.1556*.
[13] K. F. Haque, K. M. Rumman, A. Elyasi, F. Meneghello, and F. Restuccia, “MAGIC: Meta-learning adaptive gesture recognition with mmWave MIMO CSI,” in *Proc. IEEE Int. Symp. World Wirel., Mobile, Multimedia Netw.*, 2025, pp. 131–140.