



# Fusang: Graph-inspired Robust and Accurate Object Recognition on Commodity mmWave Devices

Guorong He<sup>1</sup>, Shaojie Chen<sup>1</sup>, Dan Xu<sup>1\*</sup>, Xiaojiang Chen<sup>1</sup>

Xiaohong Xie<sup>2</sup>, Xinhua Wang<sup>3</sup>, Dingyi Fang<sup>1</sup>

<sup>1</sup>Northwest University, Xi'an, China, <sup>2</sup>University at Buffalo SUNY, New York, USA, <sup>3</sup>Xidian University, Xi'an, China

guorong.nwu@gmail.com, chenshaojie@stumail.nwu.edu.cn, {xudan, xjchen, dyf}@nwu.edu.cn

xiaohongx@buffalo.edu, xinhua.wang@xidian.edu.cn

## ABSTRACT

This paper presents the design and implementation of *Fusang*, a low-barrier system that brings accurate and robust 3D object recognition to Commercial-Off-The-Shelf mmWave devices. The basic idea of *Fusang* is leveraging the large bandwidth of mmWave Radars to capture a unique set of fine-grained reflected responses generated by object shapes. Moreover, *Fusang* constructs two novel graph-structured features to robustly represent the reflected responses of the signal in the frequency domain and IQ domain, and carefully designs a neural network to accurately recognize objects even in different multipath scenarios. We have implemented a prototype of *Fusang* on a commodity mmWave Radar device. Our experiments with 24 different objects show that *Fusang* achieves a mean accuracy of 97% in different multipath environments. The code, dataset, and trained models of *Fusang* can be obtained at <https://github.com/OpenNISLab/Pro-Fusang>.

## CCS CONCEPTS

- Human-centered computing → Ubiquitous and mobile computing systems and tools;
- Computing methodologies → Feature selection; Machine learning approaches.

## KEYWORDS

Mmwave Radar, HRRP, Graph-inspired Feature, Object Recognition

### ACM Reference Format:

Guorong He<sup>1</sup>, Shaojie Chen<sup>1</sup>, Dan Xu<sup>1\*</sup>, Xiaojiang Chen<sup>1</sup> and Xiaohong Xie<sup>2</sup>, Xinhua Wang<sup>3</sup>, Dingyi Fang<sup>1</sup>. 2023. *Fusang: Graph-inspired Robust and Accurate Object Recognition on Commodity mmWave Devices*. In *The 21st Annual International Conference on Mobile Systems, Applications and Services (MobiSys '23)*, June 18–22, 2023, Helsinki, Finland. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3581791.3596849>

## 1 INTRODUCTION

**Motivation & Problem Statement.** Heading towards a connected world, object recognition as the cornerstone of perceiving the world

\*Dan Xu is the corresponding author.

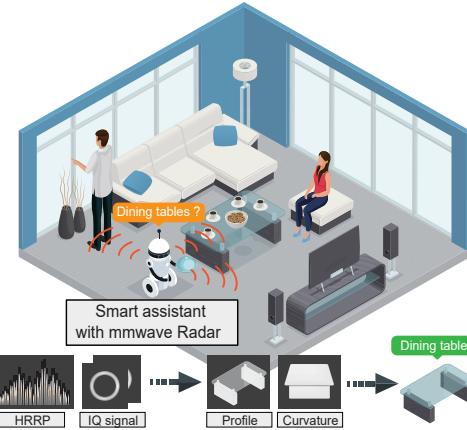
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

*MobiSys '23, June 18–22, 2023, Helsinki, Finland*

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0110-8/23/06...\$15.00

<https://doi.org/10.1145/3581791.3596849>



**Figure 1: Illustration of *Fusang*.** *Fusang* leverages the mmWave signal reflected from ambient reflections to identify 3D objects in the living room.

plays a key role in many related applications, such as autonomous driving [31], security checks [49], smart logistics [55] and assistive navigation [34, 65] for blind people or humanoid robots. However, such real-world applications often require fast and efficient discrimination of challenging and highly similar objects in complex details. For example, poor-quality object recognition can lead to serious, sometimes life-threatening accidents when the blind or self-driving cars are on the road [5, 48]. During security checks and logistics inspections, low-quality object recognition may also cause serious economic losses and safety hazards [49].

**Limitations of Prior Art.** Latest research [15, 39, 46, 63] on camera & Lidar-based 3D object recognition have achieved promising results, especially with the help of deep learning, but they suffer from severe performance degradation under adverse conditions (e.g., foggy or rainy weather, poor illumination). More significantly, optical cameras capture the scene through a true-color image representation [1, 52], making them too privacy-invasive to be implemented in private spaces, like bedroom or office. Some studies propose using WiFi [32] or RF [61] signal parameters (e.g., received signal strength, phase) to capture images of the target surface, but they are limited to capturing only a rough two-dimensional outline of the object [47]. Additionally, low-frequency signals require a larger antenna aperture [2, 38], making it challenging to deploy devices on a large scale.

Recently, object reconstruction based on mmWave devices [70, 78] has emerged as an alternative way to obtain information about 3D surroundings. By taking advantage of high-frequency wireless

signals, the fine-grained shape details of the target can be captured and modeled. The mmWave sensing systems are robust to low visibility conditions, which is suitable for some non-line-of-sight sensing tasks [42, 77]. Despite mmWave has numerous advantages, accurately recognizing targets using commercial off-the-shelf (COTS) millimeter-wave equipment remains highly challenging. The main reason is that commercial devices limit the number of antennas to control costs (just like the insect equivalent of compound eyes), resulting in low image resolution [1]. To tackle this issue, *MilliPoint* [47] suggests using Synthetic-aperture Radar (SAR) to generate 3D point cloud from mmWave radar, enabling richer information of targets. However, SAR-based methods assume that the scene must be quasistatic during Radar movement along the cross-range. Once targets move or the scanning route deviates, model reconstruction will suffer from blur and artifact. Therefore, to unlock the full potential of mmWave 3D object recognition, a high precision, low complexity, and robust system is still needed.

**Proposed solution.** This paper presents the design and development of *Fusang*<sup>1</sup>, a *low-barrier system that brings high-resolution 3D target recognition on mmWave COTS devices*. As shown in Fig. 1, a smart assistant equipped with the *Fusang* can make efficient and robust responses by relying entirely on sensor information from off-the-shelf mmWave Radar to quickly capture various object contours. Inspired by the wideband radar system [6, 66], *Fusang* directly associates the object profile with the scattering features of the radar signal by establishing the projection of the target scattering point on the line-of-sight direction *i.e.*, *high-resolution range profile* (HRRP): one-dimensional vector obtained by coherent summations of the complex time returns from target scatterers in each range cell.

Specifically, HRRP has two unique advantages compared with traditional point cloud imaging: (1) HRRP preserves more energy and structure information of target scattering along the range direction [9]. Since HRRP does not go through the traditional radar signal processing algorithm (*e.g.*, constant false alarm rate), it retains more object information than the point cloud. (2) HRRP data is easier to obtain. The HRRP data can be directly extracted by FFT of reflected signals, so this method can effectively simplify the processing flow of radar signals and is suitable for real-time target recognition tasks[76], such as blind navigation or autonomous driving compared with the SAR&ISAR.

Despite the advantages, capturing the spatial shape of a specific target from HRRP data directly is non-trivial due to the following challenges. *Firstly*, the impact of multipath in indoor environments. Although mmWave has better directionality than conventional wireless signals, the received signal at the radar antenna is still a mix of the signals reflected from both the target object and other reflectors in the environment. *Secondly*, the target’s HRRP data is inherently susceptible to different parameters and measurement conditions. Even for the same target at the same distance, the amplitude of HRRP data is obviously different, which degrades the final target recognition accuracy [9, 10]. *Thirdly*, because HRRP data is essentially the projection of the target scattering point on its distance dimension, it only preserves the relationship between

<sup>1</sup>A mythical tree described in ancient Chinese literature, it is said that it connects the sky and the earth. We named the system *Fusang*, hoping that our system can become a bridge connecting the virtual and the real world.

the coarse contour and the scattering features of the target [47, 76], lacking fine-grained information about target curvature. For some different targets with similar contours, therefore, purely relying on the HRRP data of the object is not enough to classify them.

To overcome these challenges, we first analyze the frequency spectrum energy of reflected mmWave signals through Successive Variational Mode Decomposition (SVMD). By combining this with the Radar localization function, *Fusang* is able to capture the energy envelope and multi-frequency information from the reflected signals of a target in multipath-rich environments. After frequency spectrum energy analysis, we find that even though the frequency components of the target point may have different amplitudes, its energy distribution remains stable and distinct. This enables *Fusang* to effectively mark and identify different objects.

To extract more effective features from HRRP data, *Fusang* proposes a novel *tree-like graphical structure* to represent the energy distribution of target scatterers and improve the robustness of HRRP data. Different from previous models, the proposed tree-structured features represent the dominant energy distribution of the input Radar signals with more powerful discrimination, which can be naturally and easily generalized and applied to different kinds of targets. Furthermore, to identify similarly shaped objects which cannot be distinguished by HRRP data, our key observation is that the In-phase and Quadrature (IQ) domain signal of Radar signals at different angles may have periodic changes. *Fusang* thus records the IQ samples of the reflected signals from multiple antennas at different angles, and designs the customized features in the IQ domain to achieve fine-grained curvature detection. Combining profile and curvature information, we design a customized neural network to predict the accurate 3D shape of many different objects just by using a commercial Radar.

**Summary of Results.** We have prototyped *Fusang* on an off-the-shelf mmWave Radar and conducted extensive benchmark experiments to verify its performance on data collected from 24 common objects in the at-home applications. The results of our experiments demonstrate that *Fusang* can effectively extract the profile and curvature of the object, and realize the 3D recognition of the object accurately. Specifically, in high SNR environments such as empty halls, the recognition accuracy can reach up to 99%. In the multipath-rich office, the lowest 3D object recognition accuracy for *Fusang* in our experiments is 88%. Compared to state-of-the-art HRRP-based object recognition system [9], the prediction accuracy of target of *Fusang* is improved by 1.5 $\times$  under the same SNR condition. Moreover, we also verify the practicality of *Fusang*, by examining the impact of diverse multipath environments, the distance between the objects, the angle of incidence, and the motion of radar. Results show that *Fusang* leads to a more accurate target recognition and better robustness with different influence parameters.

**Contributions.** (i) We present *Fusang*, a low-barrier system that brings high-resolution 3D target recognition, which relies solely on sensors available in an off-the-shelf mmWave Radar to capture a wide range of profile and curvature information, effectively characterizing the overall 3D object (Sec. 3.1). (ii) We design a signal reconstruction algorithm, which combines the range spectrum with SVMD to capture the multi-frequency component of the target signal. This method successfully overcomes the limitation of single

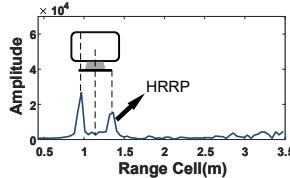


Figure 2: An HRRP sample from a target.

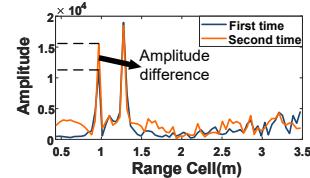


Figure 3: Same objects but different HRRPs.

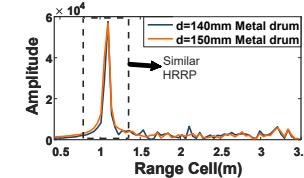


Figure 4: Different objects but same HRRPs.

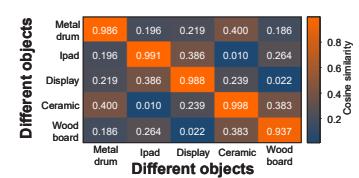


Figure 5: The similarity of HRRP energy distribution.

Radar in multipath-rich environments (Sec. 3.2). (iii) We develop a novel tree-structured model for capturing the stable energy distribution of Radar HRRP data and IQ signal to characterize the object profile and curvature robustly. In addition, *Fusang* constructs a customized hybrid network to achieve accurate recognition of the 3D target (Sec. 3.3-Sec. 3.5).

## 2 BACKGROUND AND MOTIVATION

In this section, we first introduce the generation principles of Radar HRRP data and then analyze why it can't be directly used for the 3D modeling of targets. In the third subsection, we present important observations on the HRRPs' energy distribution, which motivate the design of *Fusang*.

### 2.1 Review of Radar HRRP Data

For a typical wideband Radar, the range resolution,  $\Delta R = c/2BW$ , is far smaller than the target size, and the electromagnetic characteristic of the target can be approximated by the multiple scatterers in range cells, where  $c$  is the signal propagation speed ( $\sim 3 \times 10^8$  m/s) and  $BW$  is Radar transmit bandwidth. Accordingly, HRRP data can be considered as a projection of Radar returns from many scatterers distributed in several range bins onto the LoS, as shown in the Fig. 2. Let us consider that an object (such as the display in Fig. 2) has  $R$  reflecting points, and the scattering strength of  $r^{th}$  point is  $\sigma_r$ . The mmWave Radar transmits wide frequency-modulated continuous wave signals ( $\sim 77\text{GHz}-81\text{GHz}$ ) from each Tx antenna, which can be expressed as:

$$S_{Tx}(t) = \exp \left[ j \left( 2\pi f_c t + \pi K t^2 \right) \right], t \in [0, \tau] \quad (1)$$

where  $f_c$  and  $K$  are the starting frequency and the chirp slope of frequency-modulated continuous wave signal respectively.  $\tau$  is the chirp duration. Then the signals are reflected by the target and picked up by the Radar receiving antennas. The received signal can be expressed by a sum of all time-delayed transmitted signals:

$$S_{Rx}(t) = \sum_{r \in R} \sigma_r \cdot S_{Tx}[t - 2d_{ra}/c] \quad (2)$$

where  $d_{ra}$  represents the distance between the  $r^{th}$  reflecting point and  $a^{th}$  receiving antenna. Next, a mixer is utilized to eliminate the carrier wave in the received signal and obtain the so-called beat frequency signal:

$$S(t) = S_{Tx}(t)S_{Rx}^*(t) \approx \sum_{r \in R} \sigma_r \cdot \exp[j4\pi(f_c + Kt)d_{ra}/c] \quad (3)$$

To extract the HRRP data from  $S(t)$ , a Fast Fourier Transform (FFT) operation on the samples of  $S(t)$  within a chirp to separate the reflecting points from different range cells (See Fig. 2). Based on this operation, the distribution of target reflection points in

different positions will be reflected in the spectrum map (the two energy envelopes in Fig. 2). For each range bin, the reflected signals from the target points are projected onto the image grid to generate the two-dimensional profiles of the target, which are called HRRP data, and the HRRP samples can be defined as:

$$\begin{aligned} hrrp &= [|hrrp(1)|, |hrrp(2)|, \dots, |hrrp(i)|] \\ &= \left[ \left| \sum_{r \in R_1} \sigma_r e^{j4\pi(f_c + Kt)d_{ra}/c} \right|, \dots, \left| \sum_{r \in R_i} \sigma_r e^{j4\pi(f_c + Kt)d_{ra}/c} \right| \right] \end{aligned} \quad (4)$$

where  $|.|$  means taking absolute value and  $R_i$  denotes the target reflection points in the  $i$ -th range bin.

### 2.2 Limitation of Using HRRP Directly

For different objects, their structure differences lead to discrepancies in scatterers' magnitude (i.e.,  $\sigma_r$  with  $r \in R_i$ ), finally manifested by the distinctions of HRRP. Based on this, the HRRP can be exploited to recognize the category of unknown targets. However, achieving high-resolution 3D object recognition from COTS mmWave devices, based on HRRP data, is challenging because of two reasons.

**(1) Vulnerability of HRRP data.** The HRRP data used in practice is vulnerable to the parameters such as target distance, angle of measurement, different running speeds and measurement conditions of Radar. Even for the same object and the same antenna, the peak of each FFT bin at the same position is found to be different at consecutive moments (See Fig. 3). The sensitivity difference of the amplitude makes it difficult to directly match the reference data with the measured data, that is, the amplitude of HRRP data is difficult to be used in the identification directly.

**(2) Failure of similar objects.** Although HRRP data contains a large number of structural features of targets, it is essentially the projection of target scattering points on its distance dimension, and only achieves the correlation between the coarse contour and the scattering points of the target. Thus, the HRRP of two objects with similar two-dimensional contours will be very close and hard to distinguish. For example, Fig. 4 shows the HRRP of two metal drums with similar diameters at the same distance. It generates two similar-looking indistinguishable HRRP images, which may lead to bias or even complete invalidation of target recognition.

### 2.3 Energy Distribution of Reflected Signal

**2.3.1 HRRP Energy Distribution.** To cope with the above challenge (1), our insight is to leverage the *stability of energy distribution of target scatterers in their frequency components*. Specifically, Fig. 5 plots the cosine similarity thermal map of the HRRP data from five different objects (a total of 81920 signals) in the same scene. It is not difficult to observe that HRRP data of the same object have similar

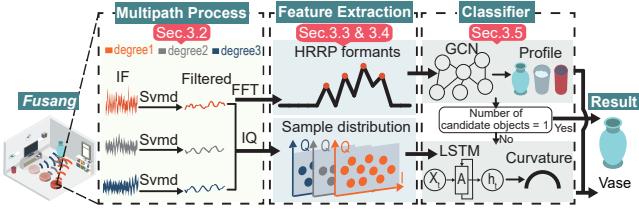


Figure 6: System overview of *Fusang*.

energy distribution in their frequency component ( $\text{corr} \geq 0.93$ ), and the energy distribution of different objects is significantly discrete ( $\text{corr} \leq 0.4$ ). We find that although the amplitude of HRRP data is dynamic and irregular, the energy distribution of HRRP data in the frequency components is highly stable and traceable. Indeed, a higher correlation between two HRRP signals implies a higher similarity between their corresponding 2D profiles. Therefore, we can correctly identify the profiles of different objects, as long as the energy distribution of HRRP is extracted.

**2.3.2 IQ Signal Distribution.** As for objects with similar shapes (e.g., challenge (2)), observing the reflected signal's spatial characteristics in the IQ domain provides a new angle to tackle this problem. Specifically, due to most objects in our daily life being axisymmetric, mainly for balance, the IQ sample distribution of the reflected signal shows periodic changes when the Radar signal illuminates the target from different spatial angles. This is because objects with different curvatures scatter electromagnetic waves to different degrees. We analyze this part in detail in Sec 3.4 and Fig. 12 plots this interesting process. Such a periodic pattern combines with the positional variation of the target surface, providing a fine-grained basis for accurately tracing the target curvatures. Based on the interesting fact between object curvature and IQ signal, we can then recognize some objects with similar shapes but different curvatures.

### 3 FUSANG DESIGN

Based on the above insights and analysis in Sec. 2.3, we design our system, *Fusang*, which achieves high-resolution and robust 3D target recognition on mmWave COTS devices through the distribution of HRRP and IQ samples. We design two novel graph-based features, namely *Branches* and *Leaves*, which are used to characterize the distribution of HRRP energy and IQ signals respectively. At the same time, a lightweight hybrid network model is designed to improve the accuracy of object recognition. Since *Fusang* directly relates the target profile features to the scattering features of Radar signals and does not need to completely illuminate the whole picture of the target object, our system can work in mobile cases with high accuracy performance compared to previous works.

#### 3.1 Overview

The overview of our system is shown in Fig. 6, which consists of four components. We introduce each component below.

- **Multipath signal processing.** Firstly, *Fusang* takes the raw samples of mmWave signals as the sole inputs. By analyzing the range spectrum of the signals, it detects the candidate range bins of targets and then generates the HRRP data of the objects. We further perform time-frequency analysis to obtain the set of

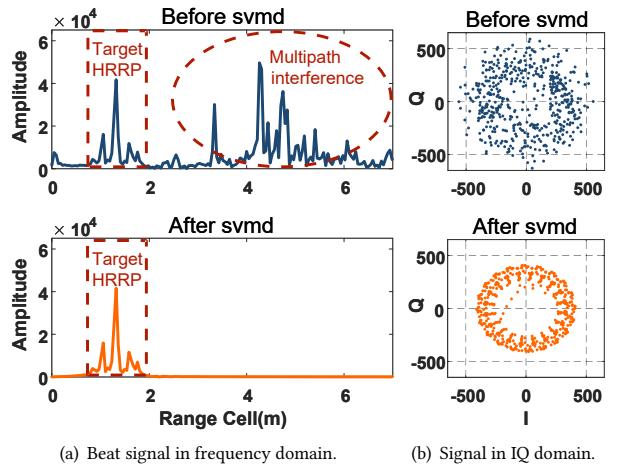


Figure 7: The results of SVMD in frequency domain (a) and IQ domain (b) for an indoor multipath scene.

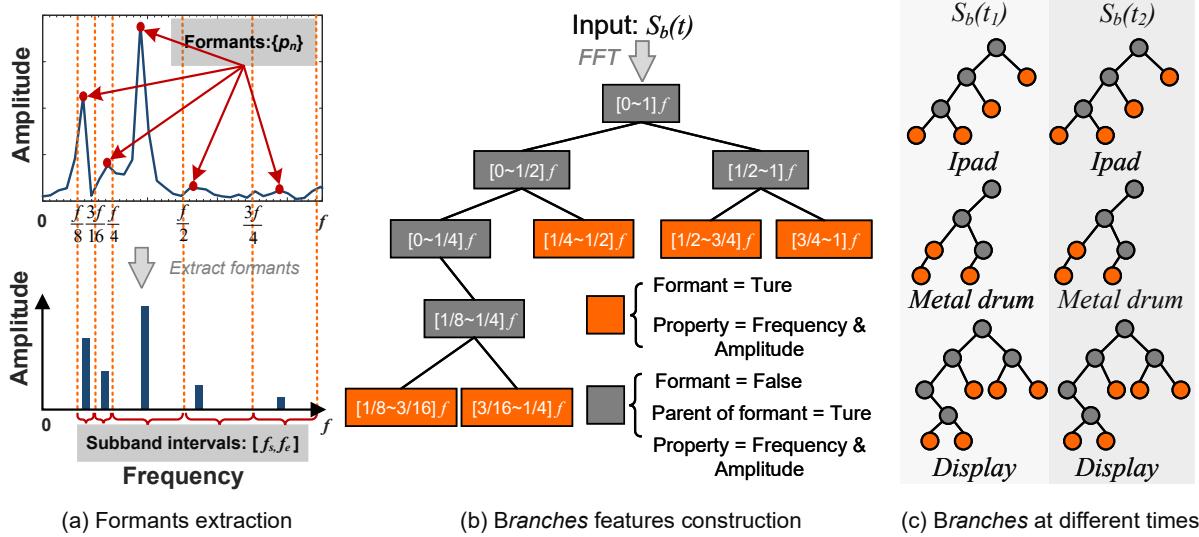
signal frequency subbands with various energy envelopes. Based on the target position on the range-angle spectrum, we can get the frequency range corresponding to the position. Then we select the appropriate frequency subbands to merge and achieve the separation of the interference signal.

- **Robust HRRP graphical representation.** After successfully screening out the target signal, based on the stability and discriminability of HRRP energy distribution in sec 2.3.1, *Fusang* constructs a graph feature based on binary tree structure, termed as *Branches*, which more efficiently and robustly extract informative representations from HRRP energy envelopes.
- **Accurate curvature feature extraction.** At the same time, in order to accurately identify the similar shape of the object, *Fusang* collects the mmWave signals reflected from multiple angles. Based on the observed periodic variation of the IQ samples at multiple angles (in the sec 2.3.2), we design a second feature, namely *Leaves*, which is used to characterize the changes of IQ samples in different angles.
- **3D object recognition.** Finally, *Fusang* takes into account both two features of the object simultaneously, and builds a customized hybrid neural network, which contains several GCN and LSTM layers to independently recognize the two-dimensional contour and curvature of the target. Further, *Fusang* concatenates all the measurements to determine the types of 3D targets.

#### 3.2 Multipath Signal Processing

To extract only the target's HRRP energy distribution from the reflected signals, *Fusang* first needs to preprocess the raw samples to remove spurious information. Multipath signal processing involves two steps: (1) Target location by range spectrum to obtain information about the target objects, including their positions and reflected signal strength, and (2) Decomposition and reconstruction of the signal by SVMD for subtracting unwanted background noise.

- 3.2.1 **Target Location.** *Fusang* first periodically searches the environment and receives the reflected signals to produce the beat signals composed of the In-phase and Quadrature-phase data. Through the beat signals, we calculate the signal strengths of all positions



**Figure 8:** Given a filtered signal, we first perform FFT and extract the set of formants (a) to represent the HRRP data energy distribution. Then, according to the formants, we construct (b) *Branches*, a binary tree, where each terminal node is a formant that falls in each frequency subband  $[f_s, f_e]$ . The tree node has two properties: the amplitude and frequency, which are used as input to the graph classifier. We construct *Branches* features for three objects with different shapes respectively, and we can see that the distribution of binary trees is significantly different in (c). In addition, *Branches* overcomes the vulnerability of HRRP data, the binary tree structure of the same object at different moments is completely consistent.

in the field of view and obtain the spectrum of the multipath environment by utilizing classic range FFT algorithms. Then, a classical CFAR algorithm [50] is implemented on the range spectrum to extract the distance information  $d_{ra}$  between the target and Radar antenna in the scene. For the target at a range  $d_{ra}$ , the resulting beat signal has a frequency  $f_b = \frac{K^2 d_{ra}}{c}$  in the signal band, where  $K$  is the slope of a frequency-modulated continuous wave. So far, we have obtained the distance between the target and the Radar and the corresponding frequency component of the target at this range.

**3.2.2 Signal Decomposition and Reconstruction.** Next, *Fusang* decomposes and reconstructs the beat signals according to the frequency component of the target to eliminate noise interference in the multipath environment. Specifically, we perform Successive Variational Mode Decomposition (SVMD) [44] to decompose the original signal into  $\kappa$  band-limited and specific frequency sub-signal  $u_\kappa$ , namely Intrinsic Mode Functions (IMF), which can be expressed as:  $S(t) = \sum_\kappa u_\kappa(t) + \theta(t)$ , where  $\theta(t)$  is the final residue, which represents the low-frequency trend of the signal  $S(t)$ . SVMD algorithm estimates the bandwidth through Gaussian smoothness of the demodulated beat signals, such as, the squared L2-norm of the gradient. Then, the constrained variational models are given by

$$\min_{\{u_\kappa\}, \{\omega_\kappa\}} \left\{ \sum_\kappa \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_\kappa(t) \right] e^{-j\omega_\kappa t} \right\|_2^2 \right\} \quad (5)$$

where  $\omega_\kappa$  is the center frequency of sub-signal  $u_\kappa$ .  $\partial_t$  represents partial derivative operation with respect to time. The symbol  $\delta$  is the Dirac distribution of sub-signal  $u_\kappa$ . The essence of SVMD is to obtain a set of mode functions and their respective center frequencies, by Gaussian smoothness, such that the band-limited mode functions reproduce the input signal exactly. Different from the

common examples of linear transform methods such as short time Fourier transform [23] and wavelet transform [3], SVMD avoids several drawbacks of linear time-frequency analysis, such as limited by Heisenberg uncertainty theorem and prior assumptions on input data [58]. As a result, SVMD is suited for non-linear and non-stationary Radar signals. The detailed theory and derivation process of SVMD can be referred to literature [17, 44].

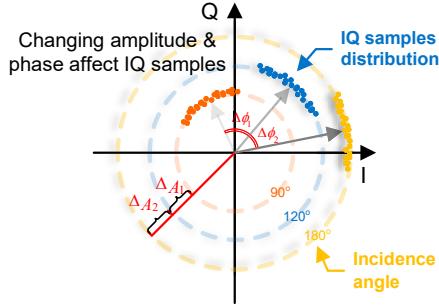
By the SVMD algorithm, we transform the original beat signal into the sum of several IMFs from high to low frequencies. Then, *Fusang* extracts the IMFs matching the target's frequency  $f_b$  and removes the high-frequency noise components. After that, we can find a group of IMFs containing the target location information, which constitutes the filtered signal  $S_b(t)$ .

$$S_b(t) = S(t) - \sum_\kappa u_\kappa(t), \quad f_b \notin \{\omega_\kappa\} \quad (6)$$

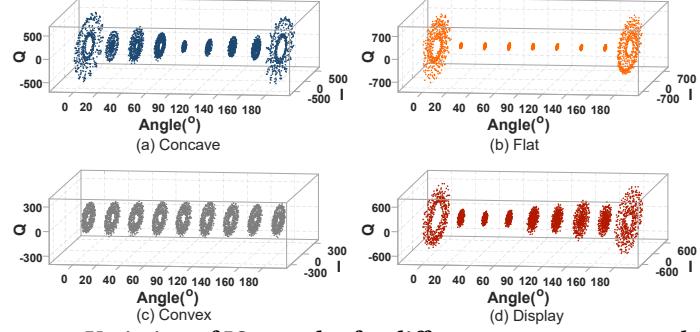
Fig. 7(a) & (b) show the signal reconstruction of our system in a common indoor multipath scenario. We can observe that *Fusang* has a positive suppression on multipath interference around the target. Such a result indicates that our method can effectively filter out the impact of multipath noise.

### 3.3 HRRP Graphical Representation

After the above steps, we extract the valid signal of the targets. However, as discussed in Sec. 2.2, direct use of HRRP data in the spectrum is not always distinguishable from each other. Therefore, in this part, we describe how *Fusang* constructs the graphical feature *Branches*, and how this feature robustly represents the energy distribution of HRRP data.



**Figure 9: Influence of surface curvature on signal IQ distribution at different angles**



**Figure 10: Variation of IQ samples for different curvatures at multiple angles.**

**3.3.1 Formant Extraction.** *Fusang* first calculates the spectrum with the FFT on the entire filtered signal  $S_b(t)$  so as to obtain the overall HRRP data of the target. Then, *Fusang* extracts the set of formants  $P = \{p_n\}$ , where  $p_n$  denotes a local maximum (i.e., formant) of the envelope of HRRP. Such formants are the direct source of the HRRP data energy distribution. Next, we iteratively bisect the frequency band until there is at most one formant that falls in each frequency subband  $[f_s, f_e]$ , as shown in Fig. 8(a), where  $f_s$  and  $f_e$  represent the starting and ending frequency of each frequency subband.

**3.3.2 Profile Feature Construction.** Based on these formants  $\{p_n\}$  and their subband intervals  $[f_s, f_e]$ , *Fusang* further constructs a full binary tree with some height, as shown in Fig. 8(b). During this process, we locate each formant in  $[f_s, f_e]$  serves as our tree node. If an interval does not contain any formant, then we remove the interval node from the binary tree. The nodes of resulted subtree cover all the frequency subbands that contain formants. With the above, *Fusang* transforms the energy distribution of different targets' HRRP into tree-structure graphs with different shapes, termed as *Branches*.

Such a graphical representation has two obvious advantages. (1) *Robust to HRRP amplitude variations*. Because HRRP data identification is a typical small-sample problem, that is, the feature dimension is high and the number of samples is small, the accuracy of parameter estimation will be affected by directly analyzing HRRP's energy distribution [19, 20], particularly under a limited amount of training data. The structure of feature *Branches* in our system is independently constructed from the formant locations of each HRRP data, which will not be affected by the amplitude variation of samples. Hence, *Branches* can more effectively extract the profile information of the object in different distances, angles, and multipath environments. (2) *No dependence on model priors*. In fact, some distributional statistical modelings can also realize the energy distribution extraction of HRRP data. However, these schemes assume that the target samples follow a certain distribution type, such as Gaussian or T-distribution [22, 59]. The selection of distribution type usually depends on the researchers' cognition of data and long-term experience accumulation, so the performance is impaired in some specific applications without prior knowledge. In this part, our feature construction does not rely on any prior knowledge and has superior feature generalization in different applications.

We construct *Branches* features from HRRP data in two consecutive time frames of different objects (See Fig. 8(c)). By comparing

Fig. 3, it's not difficult to find that even though HRRP samples have obvious jitter and dislocation, the binary tree constructed by our system is completely consistent. This is mainly because the connections between the feature nodes contain the energy distribution relationship of HRRP data, rather than directly using the amplitude and frequency of HRRP data itself. Therefore, through the *Branches*, we have a robust graphical representation of energy distribution of HRRP data. However, as discussed in Sec. 2.2, it's challenging for HRRP data to deal with two objects of similar contour. The lack of curvature information could lead to inaccurate target recognitions.

### 3.4 Curvature Feature Extraction

To deal with the above challenge, we attempt to pursue an accurate structure feature resilient to the confusion caused by similar objects. Inspired by the phenomenon that some objects can be identified via their surface curvatures [12, 43], we propose to employ the change of the object surface curvature at different angles as an auxiliary feature (discussed in Sec. 2.3). We observed an interesting fact that *the change of the object surface curvature at different angles will cause the Radar signal to exhibit different distributions in its IQ domain*, as shown in Fig. 10.

**3.4.1 Curvature Identification in Principle.** We provide a theoretical explanation for the above observations. Suppose an object is deployed in a static scenario, their effect on signal amplitude  $A$  and phase  $\phi$  is summarized in the following equations:

$$A = \frac{\epsilon P_t G_t A_{eff}}{[4\pi d_{ra}^2]^2}, \quad \Delta\phi = \frac{4\pi\Delta d_{ra}}{\lambda} \quad (7)$$

Where  $\epsilon$  is the Radar cross-section, which is a common measure of the ability of the target to reflect the Radar signal.  $P_t$ ,  $G_t$ ,  $A_{eff}$  are parameters of transmit power, antenna gain, and effective antenna area, respectively.  $\lambda$  is the wavelength of Radar signal. Because of the presence of the object, the Radar cross section will change under different angles of observation, which impact the amplitude. In addition, signal scattering caused by subtle changes in the curvature of an object's surface, will increase or decrease the phase. As illustrated in Fig. 9, clusters with different colors are the IQ signals reflected by the object when the Radar beam is incident at different angles. We have observed that the surface curvature of an object causes simultaneous changes in the amplitude and phase of the signal at different angles, resulting in varying annular distributions of IQ signals. Such difference lies in the distinction between the

curvature of different targets. Hence, we envision that collecting IQ signals from multiple angles can enhance object sensing.

We further conduct a set of experiments to test our hypothesis. we select four objects with different curvatures and detect the objects with Radar from multiple incident angles ( $0^\circ$ - $180^\circ$ ). Each time the signal is received, the object and Radar are kept at the same distance. For concave surfaces, the targets tend to collect the electromagnetic waves that hit the surface and refocus the parallel incident electromagnetic waves on the focus. Therefore, the concave surface has an obvious convergence effect on signal amplitudes compared with the flat and convex surface(Fig. 10(a)). For flat surfaces, the IQ sample distribution is strongly dependent on the angle of incidence. Due to the specular reflection, the flat plate has almost no Radar cross-section except when aligned directly toward the Radar (i.e., the angle of incidence is  $0^\circ$  or  $180^\circ$  [57]). Therefore, the IQ sample distribution of a flat object will show the extreme phenomenon of dispersing at both ends and concentrating at all other angles, as shown in Fig. 10(b). Contrary to flat surfaces, for convex surfaces in Fig. 10(c), the distribution of IQ samples is less affected by incident angle. Because the Radar cross-section of a convex surface is rarely different in angle, especially for a sphere or a cylinder [57]. As a result, their reflected signals show a nearly uniform distribution across the IQ domain. Similarly, for a complex object containing multiple curvatures, such as a display, the distribution of its reflected signal in the IQ domain will also show unique differences, see Fig. 10(d).

In conclusion, the difference in surface curvature plays a dominant role in determining the distribution of IQ signals between the two objects. In particular, most objects in our daily life are axisymmetric, the sample distribution of the reflected signal in the IQ domain will show periodic changes. Hence, the subtle changes in the surface curvature of 3D objects are closely related to the distribution of reflected signals in the IQ domain, which can be regarded as a representative feature to further identify two objects with similar shapes.

**3.4.2 Characterizing Surface Curvature.** To characterize the distribution of IQ signals at different angles, inspired by the point density statistics [13], we define a feature named *Leaves*, which can turn a large number of discrete IQ samples into a highly abstract comprehensive description.

Conceptually, *Fusang* defines a neighborhood around the center of each IQ signal, which is sufficient to enclose all the IQ sampling points of the signal. At the same time, we count the number of all IQ samples in the neighborhood, and then divide it by the area occupied by the whole cluster of signal in the neighborhood to obtain the feature *Leaves* representing the distribution of IQ samples when the incident angle is  $\theta$ , as shown in the following equation.

$$\text{Leaves}(\theta) = \frac{\text{Num}_{IQ}}{\text{area}_{\theta \in \Psi}(\alpha, \mathcal{H}^*_{A(\theta)})} \quad (8)$$

Here,  $\Psi$  is the maximum detection angle traversed by the Radar. Parameter  $\text{Num}_{IQ}$  can be obtained directly from the signal sampling points. For the area occupied by the IQ signal, we develop the following process to obtain.

**First**, to facilitate the quantification of the area occupied by signal samples in the IQ domain, we define a unit area parameter  $\alpha$ ,

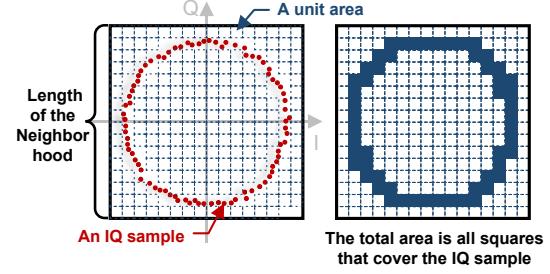


Figure 11: The statistical process of the area occupied by the IQ sampling point.

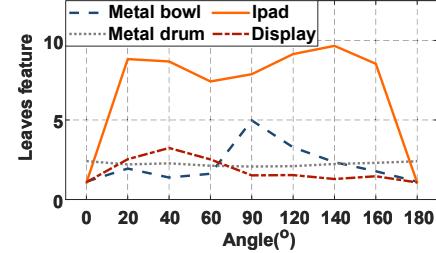


Figure 12: Variation of *Leaves* features corresponding to objects with different angles.

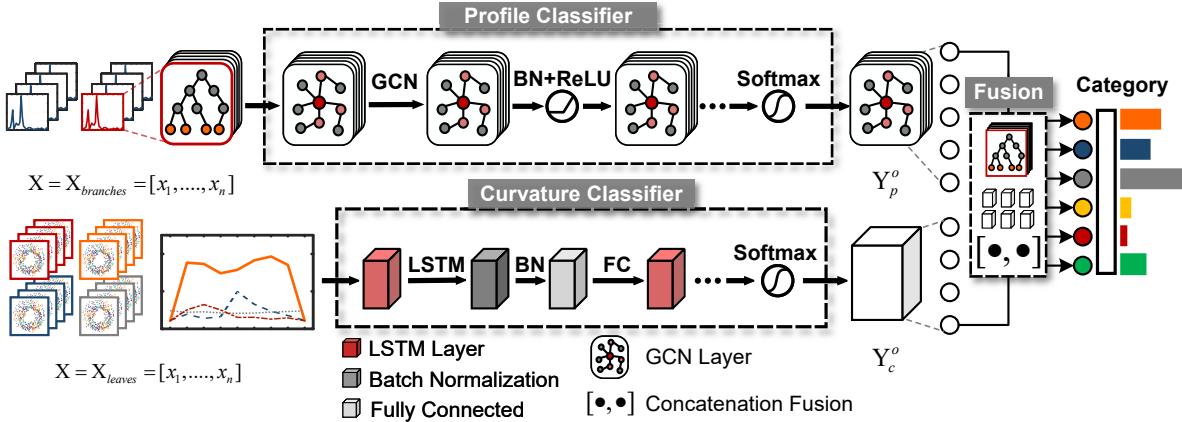
which is used to represent the measure of the minimum area in the IQ domain. Note that the value of parameter  $\alpha$  should not be too large or too small. Otherwise, the difference of *Leaves* features is not obvious. We conducted multiple experiments and determined that the optimal value for the  $\alpha$  is 20, which results in the largest difference between *Leaves* features. This value can be reused directly in other experiments without modification. **Second**, we delineate a neighborhood (See Fig. 11)  $\mathcal{H}^*_{A(\theta)}$  that near the origin of the IQ signal. To ensure coverage of the IQ signal, the length of the neighborhood must be greater than twice the maximum amplitude of the IQ signal at the current angle  $\theta$ . **Third**, according to the parameters  $\alpha$  and  $\mathcal{H}^*_{A(\theta)}$ , our system will cyclically determine which unit area each IQ sample falls in. If an IQ sample is found to be within a new unit area, add 1 to the total area. Otherwise, the total area remains the same. When the incident angle changes, the area occupied by the whole cluster of IQ signal at different angles can be obtained by repeating the above steps.

Through the above parameters, we can obtain the IQ signal distribution of objects with different curvatures at multiple angles, as shown in Fig. 12. More ‘leaves’ indicated a denser distribution of IQ samples, while fewer ‘leaves’ indicated a sparse distribution of IQ samples. Such the difference in the number of *Leaves* is used as the final curvature feature.

### 3.5 3D Object Recognition

Through these ‘branches’ and ‘leaves’, *Fusang* obtains the structured features that characterize the profiles and curvatures of an object. The final task of *Fusang* is to identify the target information contained in such signal features. To achieve this goal, we design a hybrid neural network for our system, the architecture is illustrated in Fig. 13. Before elaborating on our architecture, we first describe the problem formulation.

**3.5.1 Problem Formulation.** To perform the prediction of the profile and surface curvature of an object simultaneously, we take the



**Figure 13:** The architecture of the 3D object recognition network. Top row: the profile classifier transforms the feature vectors  $X_{\text{branches}}$  repeatedly throughout four GCN layers and then gives a prediction of the object profile. Bottom row: the curvature classifier takes the feature vectors  $X_{\text{leaves}}$  as input and applies a two-layer LSTM module to predict the curvature of object. Middle row: the fusion module takes the predicted values  $[Y_p^o, Y_c^o]$  as input and gives a comprehensive score of target category.

extracted features  $X = [X_{\text{branches}}; X_{\text{leaves}}] \in \mathbb{R}^{N_{rx} \times N_\theta}$  as inputs, where  $N_{rx}$  and  $N_\theta$  representing the number of receive antennas and the measurement angles, respectively. A labeled dataset can be denoted as  $\mathcal{D} = (X, \mathcal{Y}_p^o, \mathcal{Y}_c^o, \mathcal{Y}^o)_{o=1}^O$ , which contains  $O$  objects and  $\mathcal{Y}_p^o$  denotes the ground truth of object,  $\mathcal{Y}_p^o$  and  $\mathcal{Y}_c^o$  are the real values indicating the profile and curvature of the object label. We utilize parallel classifiers to fully leverage the high-dimensional representation of target information extracted from features  $X_{\text{branches}}$  and  $X_{\text{leaves}}$ . The outputs of these classifiers are then combined through a fusion module to obtain an accurate prediction  $\mathcal{Y}^o$  for the object. Hence, the problem we solve here is transformed into a prediction of  $\mathcal{Y}_p^o$  and  $\mathcal{Y}_c^o$  for an unlabeled input signal.

**3.5.2 Function Modules.** This part is mainly to introduce the design of our hybrid network, which consists of three modules to learn 3D object feature representations.

**Profile Classifier.** In reality, due to the target position changes, the HRRP data returned by the target may have a spectrum distortion, which leads to the deviation of the constructed tree feature  $X_{\text{branches}}$ . To achieve a distance-robust profile classifier, we need to reduce the impact of distance variation on our system. Thus, we adopt GCN as our backbone network architecture to better capture the kinetically structured property encoded in the *Branches* features. This is because the message passing of GCN can pay more attention to the connections between the formants rather than the formants themselves [24]. As shown in the top row of Fig 13, we leverage four GCN layers as the classifier for the recognition of object profile, which is provided by the Deep Graph Library [62]. Moreover, it has the ability to achieve fast convergence without requiring a large number of labeled datasets. In our experiments, the model only needs a minimum of 128 beat signals (1 frame) at each angle to achieve 80% classification accuracy. It's helpful to improve the training and testing time of the system, which may have a great advantage in some applications with few samples.

**Curvature Classifier.** As for the modeling of the curvature classifier of the target, it needs to focus on the changing features at multiple angles to restore the fine-grained surface curvature of the

object. For a changing sequence, a lot of studies have demonstrated that Long Short Term Memory (LSTM), as a variant of RNN, has an amazing ability to model long-term temporal dependencies [53]. This is suitable for us to handle the variation of curvature features  $X_{\text{leaves}}$  under multiple angles. Therefore, we adopt two common LSTM layers as the backbone of our curvature classifier, and then a Fully-connected layer (FC) to extract the overall features of the curvature prediction, as shown in the bottom row of Fig 13. In addition, at the junctions between layers, we add Batch Normalization (BN) layers to minimize overfitting.

**Fusion Module.** After carefully designing the network structures, two classifiers have learned the discriminative features of the profile and curvature respectively, and could predict these two properties of an object simultaneously. Then we design a fusion module to merge the two prediction results and give a comprehensive score of the target category. Taking a trained cup as an example, if there's nothing like a cup in the model training sets, the prediction value  $\mathcal{Y}_p^o = ['cup']$  obtained by the profile classifier will be large. So our fusion model will classify the object directly through the label of the profile classifier. If there is an object in the training sets that looks similar to the cup, the profile classifier will return two candidate predictions as '*cup*' and '*bottle*'. Then the fusion module in the middle (see Fig. 13) will concatenate the results of the two classifiers as  $['cup', \mathcal{Y}_c^o]$  and  $['bottle', \mathcal{Y}_c^o]$ . The fusion module will give the best score from the candidates according to the maximum prediction value  $\mathcal{Y}_c^o$  returned by the curvature classifier, and then obtain the final 3D object category.

## 4 IMPLEMENTATION

This section presents the system implementation, covering hardware and software platforms, data collection, and network training.

**Hardware Platform.** We implement the prototype of *Fusang* using a single-chip commercial mmWave Radar board Texas Instruments IWR1843 Boost (\$299) [29] and a data acquisition board Texas Instruments DCA1000EVM (\$499) [30]. The device contains three transmitting antennas and four receiving antennas (See Fig 14). To

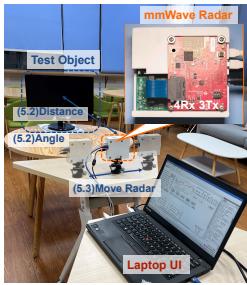


Figure 14: Implementation of system.

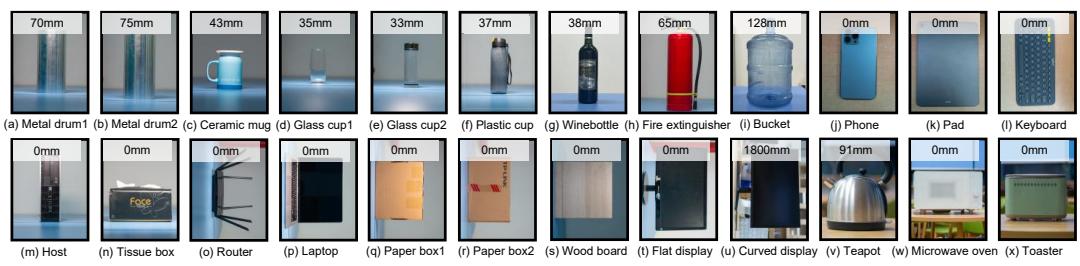


Figure 15: 24 objects used in our evaluation, which include different materials, curvatures and sizes.

ensure a minimum range resolution(3.75cm), we let one TX transmit FMCW signals starting at 77GHz with 3.99GHz bandwidth, and all RXs receive the reflected signals. Each FMCW signal is modulated by the following parameters: frequency ramp slope–20MHz/ $\mu$ s; baseband sampling rate–3Msps; number of ADC samples–512; and maximum antenna gain–10dBi.

**Software.** The software part of *Fusang* is implemented in Matlab and Python environments. We use Matlab to implement the processing of Radar signals in the early stage. In the next stage, we use Python implementation of Pytorch to build our hybrid neural network. Our software run on a computer with Intel i7-9700 CPU and Nvidia GTX 1660 GPU.

**Data Collection.** Due to the unavailability of open-source data of mmWave signal reflections about profile images and surface curvatures, we train and test *Fusang* with real data collected from our hardware platform. As shown in Fig 15, we pick 24 objects that are most frequently seen in the indoor environment (including multiple materials, curvatures and sizes) to evaluate the performance of *Fusang*, especially in offices and houses. It also includes objects with complex shapes, such as (o) a gateway with 6 antennas, (i) a keyboard with 104 keys, and (v) a teapot.

**Model Training and Testing.** For each object, we rotate each object clockwise and collect the reflected signals at 9 angles (From  $0^\circ$ - $180^\circ$ ,  $20^\circ$  at per time) spanning distances of 1-5m. It is undeniable that different perspectives and distances are needed in the training process of our system. However, not all distances and angles are necessary. In fact, our experiments have shown that the differences in graph-structured features constructed from the HRRP of the target are not significant when the target distance and angle change slightly. This is because although the frequency and amplitude of the target’s HRRP data change significantly due to changes in distance and angle, its energy distribution in the distance dimension remains stable. This is also why we choose a distance step of 1m and an angle step of  $20^\circ$ . The surface curvature of objects in Fig 15, refer to [77], labeled by the curvature radius of the object center point location. We collect a total of 14,000 frames (90GB in total), out of which 12,600 frames are used for training the network and the remaining 1,400 frames are used for testing. The diverse nature and large scale of this dataset enables us to thoroughly evaluate the robustness of *Fusang* under different environmental conditions.

**Network Parameters.** *Fusang* needs to train two independent models: GCN and LSTM. For GCN, *Fusang* explores different settings

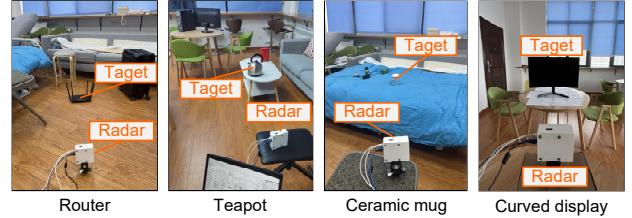


Figure 16: The setups of overall performance evaluation.

of network parameters to ensure a near-optimal model convergence. We set the initial value of model epochs to be 1000, and then, observe the loss function from GCN layers and stop its training when the model shows little to no improvement for consecutive 25 epochs. Also, we try to use different optimizers, such as, Rmsprop, Adagrad and Adam, and observe that Adam performs the best with an initial learning rate of 0.00007. As for the LSTM, we follow a similar strategy as above. Finally, we apply the RMSprop optimizer and set the learning rate as 0.0001.

## 5 EVALUATION

### 5.1 Overall Performance

**Experimental setting.** We first evaluate the overall performance of *Fusang* in a common seminar room. As shown in Fig. 16, we conduct our experiments in multiple representative indoor environments with the presence of multiple nearby interfering objects around our target. The objects are placed at varying heights across various locations in the testing area and 1m away from the Radar. To show the superiority of our object feature extraction method, we compare *Fusang* with two baselines. Baseline 1 (BS1): we directly analyze HRRP data by the method proposed in [9] instead of utilizing the graphical representations of *Branches* and *Leaves*. Baseline 2 (BS2): we only utilize the *Branches* feature to perform object recognition without curvature feature *Leaves*.

**Results.** Fig. 17 shows the confusion matrices for all 24 objects in *Fusang* and different baselines. We can observe that the average recognition accuracy of *Fusang* is 98.2%, while the average accuracy of these two baselines is 62.7% and 83.5%, respectively. Additionally, this result also shows the effectiveness of two graph-structured features of our system. Specifically, when we directly use the extracted HRRP data for target recognition (Fig 17(a)), the accuracy of object recognition is low due to the amplitude oscillation of the HRRP. By constructing binary tree feature *Branches*, we transform

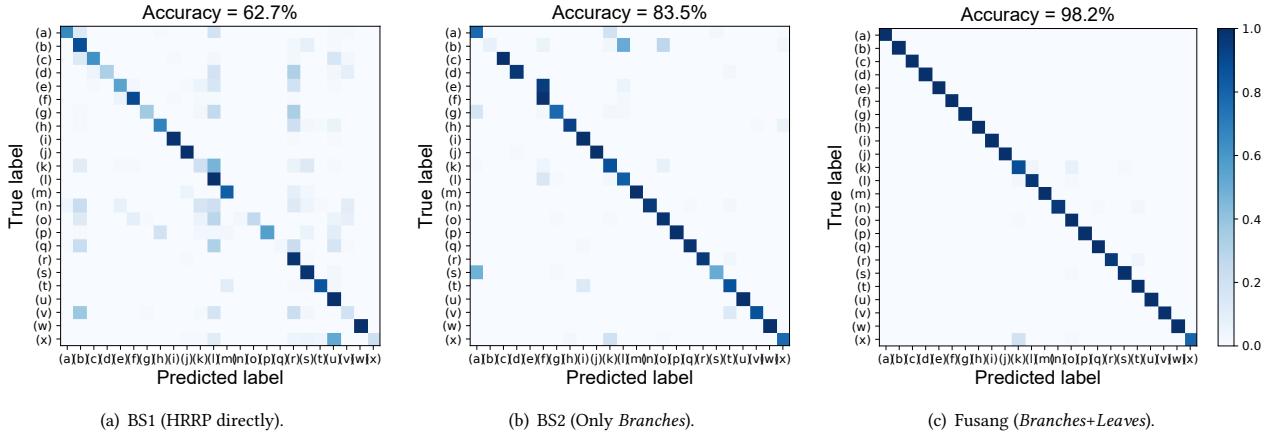


Figure 17: Confusion matrices of 24 objects for the real-world experiment.

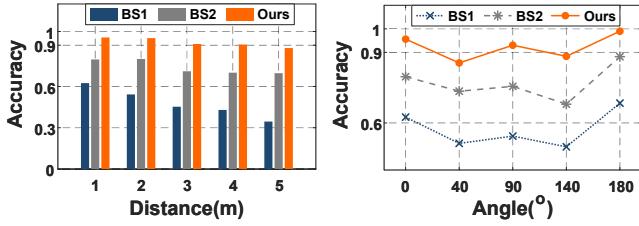


Figure 18: Accuracy under different ranges. Figure 19: Accuracy under different angles.

the energy distribution of HRRP into a robust graphical representation. Fig. 17(b) proves that feature *Branches* overcomes the influence caused by the instability of HRRP amplitude and improves target recognition accuracy by 20.8% compared with BS1. Additionally, combining Fig 17(b) and Fig. 17(c), we can also observe that using only graph-structured HRRP is not sufficient to recognize some objects with similar shapes, such as plastic cups(e) & glasses cup(f), metal drum(a) & winebottle(g), toaster(x) & pad(k). The result in Fig 17(c) shows that feature *Leaves* can effectively extract fine-grained curvature features and assist the system in achieving more accurate object recognition (improvement of 14.7% compared with BS2). In addition, it is noteworthy that our system has superior recognition performance compared to the traditional HRRP method (BS1) when identifying objects with complex shapes, with a recognition accuracy of up to 100%, such as the multi-antenna router (o), keyboard (i), and teapot (v), as shown in Fig. 16(c). This is because that complex objects have specific curvatures and shapes, resulting in more unique HRRP and IQ signal distributions. This makes tree-structured features of complex objects easier for machine learning to extract than simple objects.

## 5.2 Robustness

Robustness is a crucial issue for object recognition. So we evaluate the robustness of *Fusang* in three aspects: distance variation, angle variation, and multipath environments.

**Impact of Radar-to-object Distance.** *Experimental setting.* Note that during object sensing, *Fusang* functions as a checking spot and the target object must be within the sensing range of the radar antenna. We placed 24 objects at different distances ranging from

1m to 5m with a 1m step (covering typical indoor scenarios) for recognition. To exclude changes in HRRP data caused by different scenes and angles, objects were placed at the same incident angle and in the same room for each distance evaluation.

*Results.* Fig. 18 shows the results of our system with the variation of distance. we find although the distance variation causes a slight decrease (1.5% per meter) in accuracy, the overall recognition accuracy still remains above 89%. Meanwhile, compared with BS1 and BS2, our method improves the recognition accuracy by 2× and 1.2× respectively even at a distance of 5m. The results indicate that distance-resistant HRRP structure features are effective.

**Impact of Incidence Angle.** *Experimental setting.* Then, we evaluate how the object's incidence angle/direction with respect to the Radar affects the performance. Similar to the distance experiment, we define the default incidence angle as 0°, i.e, the Radar antenna is parallel to the surface of the object. And we keep the object at the same distance and control the incidence angle from 0° to 180°.

*Results.* Fig. 19 plots the corresponding results. We see that as the angle increases or decreases, the accuracy of system is slightly affected due to the specular reflection of objects. However, the overall average recognition accuracy of our system is still improved by 30% and 15% compared with BS1 and BS2, respectively. The lowest accuracy still reaches about 88% at 40°, revealing the angle robustness of our system.

**Impact of Different Environments.** *Experimental setting.* To validate the system's robustness under different environmental disturbances, we deploy *Fusang* in three different environments (an office, a seminar room and an empty hall), as shown in Fig. 20. For each test, objects are uniformly placed at the same distance from the radar and at the same angle of incidence to ensure the fairness of the environmental evaluation.

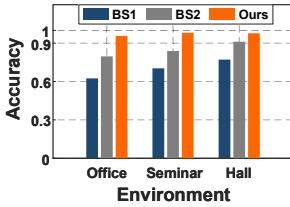
*Results.* The accuracy of object recognition in various environments is depicted in Fig. 22. We see from the figure that the average recognition accuracy of *Fusang* has a slight variation (3% when the environment changes. This is because there are many interferers in offices and seminars, and it's hard for SVMD to completely filter out the multipath influence. But, from the results, we see that the overall recognition accuracy of our system remains above 96%, proving the environment robustness of *Fusang*.



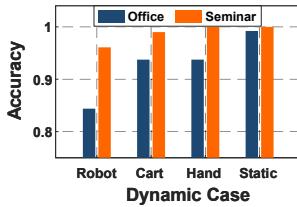
**Figure 20: Different multi-path environments.**



**Figure 21: Different dynamic case of Radar.**



**Figure 22: Accuracy under different environments.**



**Figure 23: Performance of Radar dynamic case.**

### 5.3 Dynamic Case Study

*Experimental setting.* In this section, we test our system based on some Radar dynamic cases. Our experiment only tests the situation that radar mobility, since it is more common for radar (e.g., on a Sweeping robot) to move instead of the objects. We simulate two human-computer interaction applications that require privacy protection in indoor scenarios, as shown in Fig. 21. (1) Simulate smart assistant. We mount the radar on both a handcart and a ROS-based robot, moved them at a fixed distance of 1m, and controlled the movement speed at 5m/min and 30m/min, similar to the maximum operating speed of ECOVACS-Deebot-N8 sweeping robot and Whiz sweeping robot [54]. (2) Simulate smart cane. We simulate the use of a cane by blind people to recognize objects, by randomly moving a hand-held radar in the air at an average speed of 4m/min. In three dynamic cases, we place the same test objects in front of the Radar and plot the recognition accuracy in Fig. 23.

*Results:* From the Fig. 23, it can be observed that as the speed of the radar operation increases (handheld, handcart and ROS-based robot), the accuracy of target recognition is indeed slightly affected. Specifically, in the seminar room, when the radar is stationary or moving slowly, such as when held by hand (4m/min) or pushed on a handcart (5m/min), the accuracy of target recognition is 100%. However, when the radar is mounted on a ROS-based robot and moves at a speed of 30m/min, although the recognition accuracy decreases, our system can still maintain a recognition accuracy of 96% at a distance of 1m. Similarly, in the office room, the accuracy of target recognition is 99.2% when the radar is static, 93.7% when held by hand (4m/min), 93% when mounted on a handcart (5m/min), and 84.4% when mounted on a ROS-based robot (30m/min). Through the dynamic experiment results, we observe that the *Fusang*'s object recognition performance in the office was worse than that in the seminar room. This is mainly due to the more complex multipath interference in the office compared to the seminar room, as well as the Doppler effect that occurs when objects are moving rapidly, causing the reflected signal frequency returning to the radar to shift. Consequently, the feature extraction of the binary tree graph is affected, leading to a decrease in the system's target recognition

accuracy. However, despite this, our system still ensures a recognition rate of more than 84%, even under the fast-moving radar at 30m/min, which has not been achieved in previous works [77, 78].

## 6 DISCUSSION AND LIMITATION

We mention below a few points of discussion regarding the system and scopes for improvements.

**Angle of scan.** Although the system still requires a small amount of motion radar to provide radar datasets from different angles during training, compared with SAR-based target recognition methods [47, 77], the complexity of the system is greatly reduced. Actually, after testing, our system can achieve more than 96% object recognition accuracy with only *Leaves* features on any two scanning angles. Even at 40 and 140 degrees, *Fusang* can achieve more than 87% recognition accuracy, see Sec.5 for details. We envision future deployments where intelligent reflecting surfaces or multiple radars are deployed in the environment to enrich multi-angle datasets that improve the reliability of radar target recognition.

**System delay.** We implement *Fusang* in Matlab and Python mainly demonstrate its feasibility and effectiveness. The current system is unable to process Radar data in real-time. This is mainly because the multipath signal processing is time-consuming. Specifically, on average, it takes about 0.6s to complete target location and SVMD signal reconstruction for each beat frequency signal. Nonetheless, the implementation could be optimized for real-time processing by re-implementing *Fusang* in more efficient languages, such as C++/C, and further refining the code. Meanwhile, in some environments with less multipath interference, some time-efficient signal processing algorithms like wavelet transform or empirical mode decomposition can be utilized instead of SVMD.

**Small size and unseen object.** As explained in Sec. 2.1, *Fusang* exploits the projection of the object shape on the distance dimension for robust shape sensing. Therefore, the sensing resolution of *Fusang* is limited by the range resolution of millimeter wave Radar. If the size of the target is less than the range resolution (3.75cm), their HRRP data will be sparse or even disappear. However, although the HRRP data of small objects decay rapidly at longer ranges, the signal distribution of target curvature in the IQ domain remains distinct. With the aid of the IQ signal distribution, our system achieves a remarkable average recognition accuracy of 88.96% at a testing distance of 5m. Furthermore, some super-resolution reconstruction techniques [75] can be applied to improve the system resolution.

Currently, *Fusang* cannot recognize objects that have not been learned. This is due to the lack of diverse and open-source Radar datasets for model training. One possible solution is to generate simulation data via reinforcement learning from other available sources to extend the scope of the evaluation. We will leave unseen object estimation to future exploration.

## 7 RELATED WORK

**Cameras & Lidar-based Approaches.** At present, due to the development of deep learning and optical devices, cameras and Lidar are the most widely used devices in the field of 3D object recognition and imaging [4, 26–28, 35, 41, 46, 48, 74]. Despite their ability to achieve high accuracy and real-time performance, these

technologies experience significant performance degradation in adverse conditions such as foggy weather or poor lighting. Moreover, optical cameras capture the scene through a true-color image representation [1, 52], making them privacy-invasive and unacceptable to the users to implement in their private space. In contrast, *Fusang* can achieve a more robust and privacy-friendly target perception compared with those cameras & Lidar-based technologies.

**RF-based Object Recognition and Imaging.** There are vast literature exists in the RFID, WiFi and LoRa bands perception algorithms, including object sensing [16, 52, 71], self-navigation [37] and imaging [11, 45]. *RF-ray* [16] exploits the multi-scale phase and RSS properties to estimate the shape and the material of targets simultaneously. *TagScan* [61] conducts the material estimation and target imaging utilizing the phase and RSS changes of RFID tags in a specific zone. Due to the long wavelength and limited antenna aperture, these systems require dedicated sensing devices, which are hard to provide flexible and fine-grained object sensing.

**MmWave-based Object Recognition and Imaging.** Recently, mmWave sensing, as a promising sensing technology, can achieve a very high sensing resolution, enabling various innovative applications such as object recognition and imaging [15, 47, 77, 78], human activity recognition [1, 33, 64, 68, 69], map navigation [42] and vibration sensing [60, 72]. *MilliPoint* [47] proposes 3D point cloud generation of millimeter wave Radar based on a synthetic aperture algorithm, which coherently superpositions multiple frames of Radar signals to obtain plentiful point cloud data of the target. *RSA* [78] and *Ulysses* [77] move the Radar along the trace and combine the signal parameters of specular reflections (such as RSS, AoA and AoD) to reconstruct the 2D object surface. *Geryon* [15] exploits the complementary advantages of camera and mmWave Radar to implement real-time and robust object detection. However, the above methods achieve fine-grained resolution using mechanical motion controllers, bulky arrays, or the assistance of other sensing devices. Unlike these wireless sensing systems, our work *Fusang* aims to improve the quality and efficiency of the mmWave object recognition by employing only a low-cost commodity Radar.

**HRRP-based Object Recognition.** Since HRRP contains abundant discriminative information, such as the target size and scatterer distribution, radar target recognition based on HRRP has received significant attention [7, 8, 10, 14, 18–21, 25, 36, 40, 56, 59, 73].

Early HRRP recognition methods are mainly divided into three categories: (a) statistical modeling [19, 20]; (b) feature domain transformation [21, 73]; and (c) kernel methods [7, 8]. The first type of category creates statistical models by imposing certain distributions, such as Gaussian [20], onto the HRRP. However, this approach may have limitations in terms of its ability to describe data comprehensively, the range of optimization space, and the generalization performance of the model. The second method obtains features in the transformation domain by bispectral analysis [73] and then designs proper classifiers for HRRP recognition. The over-dependency on prior knowledge, however, induces degraded performance and robustness in complex multipath environments where priors are improper. The third approach involves using kernel methods to project the HRRP data into a higher feature space. This approach can potentially result in better recognition and generalization performance, but it requires careful design of the kernels and therefore

relies heavily on the researcher’s knowledge of the data and accumulation of experience [7].

Recently, machine learning (ML) has gained significant attention for HRRP recognition, in contrast to traditional methods that heavily rely on hand-designed features. ML methods are data-driven and can automatically extract features of the HRRP such as CNNs [25, 59] and RNNs [10, 67]. The CNN can extract hierarchical spatial features from the input using cascaded convolutional and pooling layers, but it fails to capture temporal information, so it lacks robustness to HRRP data at different angles and distances [40]. To address this limitation, RNN has a sequential architecture that processes current input and historical information simultaneously to capture temporal information. However, this also introduces more noise in a complex multipath environment, which may result in limited performance [51, 56]. Moreover, due to the sensitivity of HRRP data, ML-based methods have a notable limitation. They require a significant number of precisely labeled datasets to achieve relatively accurate recognition precision [18].

The main contribution of *Fusang* compared to existing ML-HRRP solutions is that we can achieve high accuracy on object detection using much less labeled training data. We have two novel system designs to guarantee that. **First**, instead of indirectly inputting HRRP into a classifier, we propose a graphical representation of HRRP to capture the energy distribution of an object. The binary-tree graphical structure greatly helps the neural network to stably extract representative features of objects from the HRRP data. **Second**, we leverage raw IQ to capture the surface curvature features of the objects. Combining the shape-related features extracted from HRRP and the curvature-related features extracted from IQ, we can robustly classify objects with minimum training data. Our experiments demonstrate that our model still achieves more than 80% classification even if we only collect one frame of radar signals from each angle and distance as training samples.

## 8 CONCLUSIONS

*Fusang* is a novel object recognition system that only requires a single COTS mmWave Radar. It uses HRRP data and IQ samples of the reflected signals to achieve robust and accurate object recognition without relying on target augmentation or specialized optical hardware. The system employs two graph-inspired features to overcome the challenges of the vulnerability of HRRP data and the lack of accurate curvature characterization in the COTS mmWave system. With extensive experiments under 24 different objects and multipath scenarios, we have validated the promising performance of *Fusang* in accuracy and robustness.

## ACKNOWLEDGEMENTS

Thanks to the anonymous shepherd and reviewers for their valuable comments. This work is supported by National Natural Science Foundation of China under Grants (61802309, 62272388, 61972316) and the Shaanxi International Science and Technology Cooperation Program (2023-GHZD-04, 2023-GHZD-06, 2022KW-11).

## APPENDIX

The research artifact accompanying this paper is available via <https://doi.org/10.5281/zenodo.7920315>.

## REFERENCES

- [1] Aakriti Adhikari, Hem Regmi, Sanjib Sur, and Srihari Nelakuditi. 2022. MiShape: Accurate Human Silhouettes and Body Joints from Commodity Millimeter-Wave Devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 3 (2022), 1–31.
- [2] Fadel Adib, Chen-Yu Hsu, Hongzhi Mao, Dina Katabi, and Frdo Durand. 2015. Capturing the human figure through a wall. *ACM Transactions on Graphics (TOG)* 34, 6 (2015), 1–13.
- [3] A Arneodo, G Grasseau, and M Holschneider. 1988. Wavelet transform of multifractals. *Physical review letters* 61, 20 (1988), 2281.
- [4] Dejan Azinović, Ricardo Martin-Brualla, Dan B Goldman, Matthias Nießner, and Justus Thies. 2022. Neural RGB-D surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 6290–6301.
- [5] Kshitiz Bansal, Keshav Rungta, Siyuan Zhu, and Dinesh Bharadia. 2020. Pointillism: Accurate 3d bounding box estimation with multi-radars. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*. 340–353.
- [6] William W Camp, Joseph T Mayhan, and Robert M O'Donnell. 2000. Wideband radar for ballistic missile defense and range-Doppler imaging of satellites. *Lincoln Laboratory Journal* 12, 2 (2000), 267–280.
- [7] Bo Chen, Hongwei Liu, and Zheng Bao. 2008. A kernel optimization method based on the localized kernel Fisher criterion. *Pattern Recognition* 41, 3 (2008), 1098–1109.
- [8] Bo Chen, Li Yuan, Hongwei Liu, and Zheng Bao. 2007. Kernel subclass discriminant analysis. *Neurocomputing* 71, 1–3 (2007), 455–458.
- [9] Jian Chen, Lan Du, Guanbo Guo, Linwei Yin, and Di Wei. 2022. Target-attentional CNN for Radar Automatic Target Recognition with HRRP. *Signal Processing* 196 (2022), 108497.
- [10] Wenchao Chen, Bo Chen, Xiaojun Peng, Jiaqi Liu, Yang Yang, Hao Zhang, and Hongwei Liu. 2021. Tensor RNN with Bayesian nonparametric mixture for radar HRRP modeling and target recognition. *IEEE Transactions on Signal Processing* 69 (2021), 1995–2009.
- [11] Zhe Chen, Tianyue Zheng, and Jun Luo. 2021. Octopus: a practical and versatile wideband MIMO sensing platform. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*. 601–614.
- [12] Yayun Cheng, Yingxin Wang, Yingying Niu, Harvey Rutt, and Ziran Zhao. 2021. Physically Based Object Contour Edge Display Using Adjustable Linear Polarization Ratio for Passive Millimeter-Wave Security Imaging. *IEEE Transactions on Geoscience and Remote Sensing* 59, 4 (2021), 3177–3191.
- [13] Dorin Comaniciu and Peter Meer. 2002. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on pattern analysis and machine intelligence* 24, 5 (2002), 603–619.
- [14] Keith Copsey and Andrew Webb. 2003. Bayesian gamma mixture model approach to radar target recognition. *IEEE Transactions on Aerospace and Electronic systems* 39, 4 (2003), 1201–1217.
- [15] Kaikai Deng, Dong Zhao, Qiaoyue Han, Shuyue Wang, Zihan Zhang, Anfu Zhou, and Huadong Ma. 2022. Geryon: Edge Assisted Real-time and Robust Object Detection on Drones via mmWave Radar and Camera Fusion. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 3 (2022), 1–27.
- [16] Han Ding, Linwei Zhai, Cui Zhao, Songjiang Hou, Ge Wang, Wei Xi, Jizhong Zhao, and Yihong Gong. 2021. RF-ray: Joint RF and Linguistics Domain Learning for Object Recognition. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–24.
- [17] Konstantin Dragomiretskiy and Dominique Zosso. 2013. Variational mode decomposition. *IEEE transactions on signal processing* 62, 3 (2013), 531–544.
- [18] Chuan Du, Yulai Cong, Lei Zhang, Dandan Guo, and Song Wei. 2022. A Practical Deceptive Jamming Method Based on Vulnerable Location Awareness Adversarial Attack for Radar HRRP Target Recognition. *IEEE Transactions on Information Forensics and Security* 17 (2022), 2410–2424.
- [19] Lan Du, Jian Chen, Jing Hu, Yang Li, and Hua He. 2019. Statistical modeling with label constraint for radar target recognition. *IEEE Trans. Aerospace Electron. Systems* 56, 2 (2019), 1026–1044.
- [20] Lan Du, Hongwei Liu, and Zheng Bao. 2008. Radar HRRP statistical recognition: parametric model and model selection. *IEEE Transactions on Signal Processing* 56, 5 (2008), 1931–1944.
- [21] Lan Du, Hongwei Liu, Zheng Bao, and Mengdao Xing. 2005. Radar HRRP target recognition based on higher order spectra. *IEEE Transactions on Signal Processing* 53, 7 (2005), 2359–2368.
- [22] Lan Du, Hongwei Liu, Zheng Bao, and Junying Zhang. 2006. A two-distribution compounded statistical model for radar HRRP target recognition. *IEEE Transactions on Signal Processing* 54, 6 (2006), 2226–2238.
- [23] Lutfiye Durak and Orhan Arikán. 2003. Short-time Fourier transform: two fundamental properties and an optimal implementation. *IEEE Transactions on Signal Processing* 51, 5 (2003), 1231–1242.
- [24] Vijay Prakash Dwivedi, Chaitanya K Joshi, Thomas Laurent, Yoshua Bengio, and Xavier Bresson. 2020. Benchmarking graph neural networks. *arXiv preprint arXiv:2003.00982* (2020).
- [25] Bo Feng, Bo Chen, and Hongwei Liu. 2017. Radar HRRP target recognition with deep networks. *Pattern Recognition* 61 (2017), 379–393.
- [26] Yongjie Guan, Xueyu Hou, Nan Wu, Bo Han, and Tao Han. 2022. DeepMix: mobility-aware, lightweight, and hybrid 3D object detection for headsets. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*. 28–41.
- [27] Jordan SK Hu, Tianshu Kuai, and Steven L Waslander. 2022. Point density-aware voxels for lidar 3d object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 8469–8478.
- [28] Kuan-Chih Huang, Tsung-Han Wu, Hung-Ting Su, and Winston H Hsu. 2022. MonoDTR: Monocular 3D Object Detection with Depth-Aware Transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 4012–4021.
- [29] Texas Instruments Incorporated. 2020. IWR1843: Single-chip 76-GHz to 81-GHz industrial radar sensor integrating DSP, MCU and radar accelerator. <https://www.ti.com/product/IWR1843>.
- [30] Texas Instruments Incorporated. 2020. Real-time data-capture adapter for radar sensing evaluation module. <http://www.ti.com/tool/DCA1000EVM>.
- [31] Shiqi Jiang, Zhiqi Lin, Yuanchun Li, Yuanchao Shu, and Yunxin Liu. 2021. Flexible high-resolution object detection on edge devices with tunable latency. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*. 559–572.
- [32] Mohammad Hadi Kefayati, Vahid Pourahmadi, and Hassan Aghaeinia. 2022. Multi-view WiFi imaging. *Signal Processing* 197 (2022), 108552.
- [33] Hao Kong, Xiangyu Xu, Jiadi Yu, Qilin Chen, Chenguang Ma, Yingying Chen, Yi-Chao Chen, and Linghe Kong. 2022. m3Track: mmwave-based multi-user 3D posture tracking. In *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*. 491–503.
- [34] Bing Li, Juan Pablo Munoz, Xuejian Rong, Qingtian Chen, Jizhong Xiao, Yingli Tian, Aries Arditi, and Mohammed Yousuf. 2018. Vision-based mobile indoor assistive navigation aid for blind people. *IEEE transactions on mobile computing* 18, 3 (2018), 702–714.
- [35] Yingwei Li, Adams Wei Yu, Tianjian Meng, Ben Caine, Jiquan Ngiam, Daiyi Peng, Junyang Shen, Yifeng Lu, Denny Zhou, Quoc V Le, et al. 2022. Deepfusion: Lidar-camera deep fusion for multi-modal 3d object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 17182–17191.
- [36] Leiyao Liao, Lan Du, and Jian Chen. 2021. Class factorized complex variational auto-encoder for HRR radar target recognition. *Signal Processing* 182 (2021), 107932.
- [37] Jun Liu, Jiayao Gao, Sanjay Jha, and Wen Hu. 2021. Seiros: leveraging multiple channels for LoRaWAN indoor and outdoor localization. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*. 656–669.
- [38] Luyang Liu, Hongyu Li, and Marco Gruteser. 2019. Edge assisted real-time object detection for mobile augmented reality. In *The 25th annual international conference on mobile computing and networking*. 1–16.
- [39] Xiulong Liu, Dongdong Liu, Jiwu Zhang, Tao Gu, and Kepiu Li. 2021. RFID and camera fusion for recognition of human-object interactions. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*. 296–308.
- [40] Xiaodan Liu, Li Wang, and Xueru Bai. 2022. End-to-End Radar HRRP Target Recognition Based on Integrated Denoising and Recognition Network. *Remote Sensing* 14, 20 (2022), 5254.
- [41] Ziwei Liu, Tianyue Zheng, Chao Hu, Yanbing Yang, Yimao Sun, Yi Zhang, Zhe Chen, Liangyu Chen, and Jun Luo. 2022. CORE-lens: simultaneous communication and object recognition with disentangled-GAN cameras. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 172–185.
- [42] Chris Xiaoxuan Lu, Stefano Rosa, Peijun Zhao, Bing Wang, Changhao Chen, John A. Stankovic, Niki Trigoni, and Andrew Markham. 2020. See through Smoke: Robust Indoor Mapping with Low-Cost MmWave Radar. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services* (Toronto, Ontario, Canada) (*MobiSys '20*). Association for Computing Machinery, New York, NY, USA, 14–27. <https://doi.org/10.1145/3386901.3388945>
- [43] A. A. M. Muzaahid, Wanggen Wan, Ferdous Sohel, Lianyao Wu, and Li Hou. 2021. CurveNet: Curvature-Based Multitask Learning Deep Networks for 3D Object Recognition. *IEEE/CAA Journal of Automatica Sinica* 8, 6 (2021), 1177–1187.
- [44] Mojtaba Nazari and Sayed Mahmoud Sakhaei. 2020. Successive variational mode decomposition. *Signal Processing* 174 (2020), 107610.
- [45] Anurag Pallaprolu, Belal Korany, and Yasamin Mostofi. 2022. Wiffract: a new foundation for RF imaging via edge tracing. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 255–267.
- [46] Xuran Pan, Zhuofan Xia, Shiji Song, Li Erran Li, and Gao Huang. 2021. 3d object detection with pointformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7463–7472.
- [47] Kun Qian, Zhaoyuan He, and Xinyu Zhang. 2020. 3D point cloud generation with millimeter-wave radar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–23.

- [48] Rui Qian, Xin Lai, and Xirong Li. 2022. 3D object detection for autonomous driving: a survey. *Pattern Recognition* (2022), 108796.
- [49] Hem Regmi, Moh Sabbir Saadat, Sanjib Sur, and Srihari Nelakuditi. 2021. SquiggleMilli: Approximating SAR Imaging on Mobile Millimeter-Wave Devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–26.
- [50] Hermann Rohling. 1983. Radar CFAR thresholding in clutter and multiple target situations. *IEEE transactions on aerospace and electronic systems* 4 (1983), 608–621.
- [51] Dhrubojoyoti Roy, Sangeeta Srivastava, Aditya Kusupati, Pranshu Jain, Manik Varma, and Anish Arora. 2021. One size does not fit all: Multi-scale, cascaded RNNs for radar classification. *ACM Transactions on Sensor Networks (TOSN)* 17, 2 (2021), 1–27.
- [52] Meng Shen, Yaqian Wei, Zelin Liao, and Liehuang Zhu. 2021. IriTrack: Face presentation attack detection using iris tracking. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 2 (2021), 1–21.
- [53] Chenyang Si, Wentao Chen, Wei Wang, Liang Wang, and Tieniu Tan. 2019. An attention enhanced graph convolutional lstm network for skeleton-based action recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 1227–1236.
- [54] SoftBank. [n.d.]. Whiz robot. <https://apac.softbankrobotics.com/apac/robots/>.
- [55] Yanxing Song, F Richard Yu, Li Zhou, Xi Yang, and Zefang He. 2020. Applications of the Internet of things (IoT) in smart logistics: a comprehensive survey. *IEEE Internet of Things Journal* 8, 6 (2020), 4250–4274.
- [56] Yiheng Song, Qiang Zhou, Wei Yang, Yanhua Wang, Cheng Hu, and Xueyao Hu. 2022. Multi-View HRRP Generation With Aspect-Directed Attention GAN. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 15 (2022), 7643–7656.
- [57] Kenneth S Suslick. 2001. Encyclopedia of physical science and technology. *Sono-luminescence and sonochemistry*, 3rd edn. Elsevier Science Ltd, Massachusetts (2001), 1–20.
- [58] Naveed ur Rehman and Hania Aftab. 2019. Multivariate variational mode decomposition. *IEEE Transactions on Signal Processing* 67, 23 (2019), 6039–6052.
- [59] Jinwei Wan, Bo Chen, Bin Xu, Hongwei Liu, and Lin Jin. 2019. Convolutional neural networks for radar HRRP target recognition and rejection. *EURASIP Journal on Advances in Signal Processing* 2019, 1 (2019), 1–17.
- [60] Chao Wang, Feng Lin, Tiantian Liu, Kaidi Zheng, Zhibo Wang, Zhengxiong Li, Ming-Chun Huang, Wenyao Xu, and Kui Ren. 2022. mmEve: eavesdropping on smartphone's earpiece via COTS mmWave device. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 338–351.
- [61] Ju Wang, Jie Xiong, Xiaojiang Chen, Hongbo Jiang, Rajesh Krishna Balan, and Dingyi Fang. 2017. TagScan: Simultaneous target imaging and material identification with commodity RFID devices. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*. 288–300.
- [62] Minji Wang, Da Zheng, Zihao Ye, Quan Gan, Mufei Li, Xiang Song, Jinjing Zhou, Chao Ma, Lingfan Yu, Yu Gai, Tianjun Xiao, Tong He, George Karypis, Jinyang Li, and Zheng Zhang. 2019. Deep Graph Library: A Graph-Centric, Highly-Performant Package for Graph Neural Networks. *arXiv preprint arXiv:1909.01315* (2019).
- [63] Peng-Shuai Wang, Yang Liu, and Xin Tong. 2022. Dual Octree Graph Networks for Learning Adaptive Volumetric Shape Representations. *ACM Trans. Graph.* 41, 4, Article 103 (Jul 2022), 15 pages. <https://doi.org/10.1145/3528223.3530087>
- [64] Haowen Wei, Ziheng Li, Alexander D Galvan, Zhioran Su, Xiao Zhang, Kaveh Pahlavan, and Erin T Solovey. 2022. IndexPen: Two-Finger Text Input with Millimeter-Wave Radar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–39.
- [65] Jiaqing Xiong, Jian Chen, and Pooi See Lee. 2021. Functional fibers and fabrics for soft robotics, wearables, and human–robot interface. *Advanced Materials* 33, 19 (2021), 2002640.
- [66] Xiangyu Xiong, Zhenmiao Deng, Wei Qi, Hai Ou, and Zijian Cui. 2018. A novel high-precision range estimation method based on phase of wideband radar echo. *IEEE Transactions on Geoscience and Remote Sensing* 57, 6 (2018), 3392–3403.
- [67] Bin Xu, Bo Chen, Jinwei Wan, Hongwei Liu, and Lin Jin. 2019. Target-aware recurrent attentional network for radar HRRP target recognition. *Signal Processing* 155 (2019), 268–280.
- [68] Weiyi Xu, Wenfan Song, Jianwei Liu, Yajie Liu, Xin Cui, Yuanqing Zheng, Jinsong Han, Xinhui Wang, and Kui Ren. 2022. Mask does not matter: anti-spoofing face authentication using mmWave without on-site registration. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 310–323.
- [69] Hongfei Xue, Yan Ju, Chenglin Miao, Yijiang Wang, Shiyang Wang, Aidong Zhang, and Lu Su. 2021. mmMesh: Towards 3D real-time dynamic human mesh construction using millimeter-wave. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*. 269–282.
- [70] Jun Yu, Xinlong Hao, Xinjian Gao, Qiang Sun, Yuyu Liu, Peng Chang, Zhong Zhang, Fang Gao, and Feng Shuang. 2021. Radar Object Detection Using Data Merging, Enhancement and Fusion. In *Proceedings of the 2021 International Conference on Multimedia Retrieval*. 566–572.
- [71] Fusang Zhang, Zhaoxin Chang, Kai Niu, Jie Xiong, Beihong Jin, Qin Lv, and Daqing Zhang. 2020. Exploring lora for long-range through-wall sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–27.
- [72] Jia Zhang, Yinian Zhou, Rui Xi, Shuai Li, Junchen Guo, and Yuan He. 2022. AmbiEar: mmWave Based Voice Recognition in NLoS Scenarios. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 3 (2022), 1–25.
- [73] Xian-Da Zhang, Yu Shi, and Zheng Bao. 2001. A new feature vector using selected bispectra for signal classification with application in radar target recognition. *IEEE Transactions on Signal Processing* 49, 9 (2001), 1875–1885.
- [74] Yifan Zhang, Qingyong Hu, Guoquan Xu, Yanxin Ma, Jianwei Wan, and Yulan Guo. 2022. Not all points are equal: Learning highly efficient point-based detectors for 3d lidar point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 18953–18962.
- [75] Le Zheng and Xiaodong Wang. 2017. Super-resolution delay-Doppler estimation for OFDM passive radar. *IEEE Transactions on Signal Processing* 65, 9 (2017), 2197–2210.
- [76] Liyang Zhu, Gang Xiong, and Wenxian Yu. 2018. Radar HRRP group-target recognition based on combined methods in the background of sea clutter. In *2018 International Conference on Radar (RADAR)*. IEEE, 1–6.
- [77] Yanzi Zhu, Yuanshun Yao, Ben Y Zhao, and Haitao Zheng. 2017. Object recognition and navigation using a single networking device. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. 265–277.
- [78] Yanzi Zhu, Yibo Zhu, Ben Y Zhao, and Haitao Zheng. 2015. Reusing 60ghz radios for mobile radar imaging. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. 103–116.