

NeRF-enabled Analysis-Through-Synthesis for ISAR Imaging of Small Everyday Objects with Sparse and Noisy UWB Radar Data

Md Farhan Tasnim Oshim¹, Albert Reed², Suren Jayasuriya³, and Tauhidur Rahman⁴

Abstract—Inverse Synthetic Aperture Radar (ISAR) imaging presents a formidable challenge when it comes to small everyday objects due to their limited Radar Cross-Section (RCS) and the inherent resolution constraints of radar systems. Existing ISAR reconstruction methods including backprojection (BP) often require complex setups and controlled environments, rendering them impractical for many real-world noisy scenarios. In this paper, we propose a novel Analysis-through-Synthesis (ATS) framework enabled by Neural Radiance Fields (NeRF) for high-resolution coherent ISAR imaging of small objects using sparse and noisy Ultra-Wideband (UWB) radar data with an inexpensive and portable setup. Our end-to-end framework integrates ultra-wideband radar wave propagation, reflection characteristics, and scene priors, enabling efficient 2D scene reconstruction without the need for costly anechoic chambers or complex measurement test beds. With qualitative and quantitative comparisons, we demonstrate that the proposed method outperforms traditional techniques and generates ISAR images of complex scenes with multiple targets and complex structures in Non-Line-of-Sight (NLOS) and noisy scenarios, particularly with limited number of views and sparse UWB radar scans. This work represents a significant step towards practical, cost-effective ISAR imaging of small everyday objects, with broad implications for robotics and mobile sensing applications.

I. INTRODUCTION

Inverse Synthetic Aperture Radar (ISAR) imaging is a standard radar mode employed for target identification, typically involving a stationary radar and a maneuvering target to collect various aspects of target reflectivity through motion. To achieve a high-resolution image of the target in the scene, reflected signals are coherently accumulated to obtain a sufficiently large synthetic aperture. Although Synthetic Aperture Radar (SAR) imaging of large objects such as UAVs, drones, vessels, buildings, and cities is widespread [1], [2], [3], imaging of small targets remains challenging due to their small Radar Cross-Section (RCS) and limitations in radar range resolution. Existing ISAR image reconstruction methods typically require complex hardware setups [4], high-precision measurement test beds [5], or noise-less expensive

anechoic chambers [6], [7]. We propose an analysis-through-synthesis optimization framework that leverages recent advances in neural rendering [8] to perform high-resolution radar imaging that works in diverse conditions employing merely an inexpensive, easy-to-use, portable setup. The proposed framework integrates the physics governing radar signal formation, scene priors, and noise models to enable coherent ISAR image reconstructions.

The Neural Radiance Field (NeRF) technique [8] employs neural networks and differentiable volume rendering to generate new perspectives of 3D scenes from 2D images. Since its inception, NeRF has spurred substantial research efforts [9], [10], resulting in optimizations and expansions that improve accuracy and efficiency in synthesizing novel views and reconstructing 3D geometry from measurements. Recently, these techniques have extended to other sensing modalities such as sonar [11], [12], [13], radar [14], [15], [16], [17], and lidar [18], [19], [20]. Reconstructing large targets such as satellites, buildings, or city maps has been explored using NeRF for these modalities. While RaNeRF [15] and ISAR-NeRF [16] focus on neural rendering for space target reconstruction, CSAR-Incoherent [17] and MOISST [19] reconstruct urban scenes. However, imaging small targets has always posed a challenge due to their minimal Radar Cross-Section (RCS) and sensor resolution limitations. Recent work of Reed et al. [11] tackles small target synthesis with sonar but faces challenges with imaging Non-Line-of-Sight (NLOS) scenes due to the acoustic signal's inability to penetrate through different materials.

We present a novel 2D ISAR reconstruction algorithm, utilizing an analysis-through-synthesis optimization approach enabled by conventional NeRF [8] techniques, yet diverging notably in sampling (spherical instead of line sampling) and output representation (time series rather than intensity). Through numerous experiments on simulated and hardware-measured data, we demonstrate both quantitative and qualitative superiority of our approach over traditional techniques. The key contributions of our work can be summarized as -

- We introduce an end-to-end analysis-through-synthesis (ATS) framework that integrates ultra-wideband radar wave propagation and reflection characteristics to facilitate 2D scene reconstruction.
- Our approach generates ISAR images without any costly anechoic chambers or complex measurement test beds, thereby reducing both cost and computation time in the reconstruction process.
- Our method demonstrates superior performance compared to conventional backprojection (BP) in both sim-

*This work was partially supported by ONR grant N00014-23-1-2406, NSF CCF-2326905, Raytheon, Inc., Qualcomm, Google, Halicioğlu Data Science Institute – UC San Diego, and Manning College of Information & Computer Sciences – UMass Amherst.

¹Md Farhan Tasnim Oshim is with Manning College of Information and Computer Sciences, University of Massachusetts Amherst, MA, USA farhanoshim@cs.umass.edu

²Albert Reed with the Kitware INC, NC, USA albertnm123@gmail.com

³Suren Jayasuriya is with the Schools of Arts, Media and Engineering and Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ, USA sjayasur@asu.edu

⁴Tauhidur Rahman is with Halicioğlu Data Science Institute, University of California San Diego, CA, USA trahman@ucsd.edu

ulated and real scenes with multiple targets and complex structures in Non-Line-of-Sight (NLOS) and noisy environments, particularly with limited number of views and sparse UWB radar scans.

II. RELATED WORKS

A. ISAR Imaging

The concept of inverse synthetic aperture radar (ISAR) emerged to resolve imaging scenarios involving moving targets and stationary radars. Circular ISAR methods, specifically designed for stationary radar systems, facilitate high-resolution imaging of rotating targets by leveraging their circular motion. In ISAR imaging, conventional reconstruction methods often necessitate intricate hardware configurations [4], precise measurement environments [5], or costly anechoic chambers with minimal electromagnetic interference [6], [7]. These requirements pose challenges in terms of accessibility and resource expenditure, limiting the widespread adoption and scalability. Common approaches for ISAR imaging include time-domain backprojection [23], a technique that reconstructs images by correlating received radar data with expected echoes from target positions across the scene, the Range-Doppler technique [24], which processes radar data collected over multiple rotations, and the Polar Format Algorithm (PFA) [25], which directly transforms radar data from polar to Cartesian coordinates for wavefront curvature correction to compensate for distortions in received signals. However, these algorithms often face challenges of non-uniform motion, cluttered environments, limited measurement coverage, noise, motion compensation errors, and limited resolution.

B. Deep Learning-based Imaging

Recent advancements in radar signal processing, coupled with the emergence of machine learning techniques, have

shown promise in overcoming these limitations. Deep learning, in particular, has emerged as a powerful tool for radar image enhancement, target recognition, and clutter suppression. Convolutional neural networks (CNNs) and generative adversarial networks (GANs) have demonstrated remarkable capabilities in radar image processing. For ISAR imaging, Hu et al. [26] introduced a CNN-based approach for reconstructing high-resolution images from radar measurements. Additionally, Tan et al. [27] demonstrated a CNN-based method for SAR image denoising, showcasing the potential of CNNs in improving SAR image quality and utility. GANs, on the other hand, have been used for radar image enhancement, exploiting their ability to generate realistic and high-fidelity reconstructions from noisy or degraded input [28], [29].

C. NeRF-based Techniques

The Neural Radiance Fields (NeRF) framework [8], initially introduced for synthesizing photorealistic 3D scenes from images, presents a fascinating approach to improve ISAR imaging capabilities. Capitalizing on NeRF's ability to generate novel views, researchers have adopted this approach to tackle challenges in ISAR imaging including resolution enhancement, motion compensation, and clutter suppression exemplified in CoIR [14], RaNeRF [15], ISAR-NeRF [16], and CSAR-Incoherent [17] works. These techniques have explored reconstructing large targets like satellites, buildings, or even city maps using NeRF. Notably, RaNeRF [15] and ISAR-NeRF [16] focus on neural rendering for space target reconstruction, while CoIR [14], CSAR-Incoherent [17], and MOISST [19] target the reconstruction of outdoor and urban scenes. However, these approaches mostly improve on already reconstructed SAR images for novel view synthesis rather than incorporating radar measurements as input signals themselves. Our approach, on the other hand, employs an analysis-through-synthesis optimization method, using an implicit neural representation akin to NeRF, to predict point scatterers within the scene and synthesize radar measurements through a differentiable forward model, optimizing the network by minimizing the loss between synthesized and actual data.

III. HARDWARE AND MEASUREMENT SETUP

We utilize a monostatic time domain Ultra-WideBand (UWB) Impulse Radar P440 [30] with time windowing capabilities for sensing, allowing us to exclude unwanted signal reflections and conduct indoor measurements without costly anechoic chambers. It operates from 3.1 to 4.8 GHz frequency centering at 4.3 GHz. Figure 1 illustrates the radar hardware and the measurement setup. Our ISAR imaging measurement in an ordinary room employs the UWB radar positioned at a fixed location, using coherent pulse integration to enhance the signal-to-noise ratio, with the radar sampling rate exceeding the maximum Doppler extent for coherent processing into unaliased ISAR image. The procedure involves placing the object of interest on a rotating table in front of the radar system. We leverage JAYEGT [21],

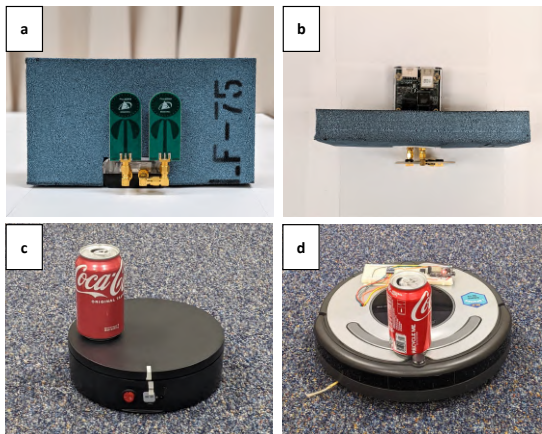


Fig. 1: Hardware and Measurement Setup: (a) Front view of Time Domain P440 radar module with absorber behind the antenna (b) Top view of the radar (c) JAYEGT [21] electronically motorized turn-table (d) Electronically enhanced Floor Cleaning Robot [22]

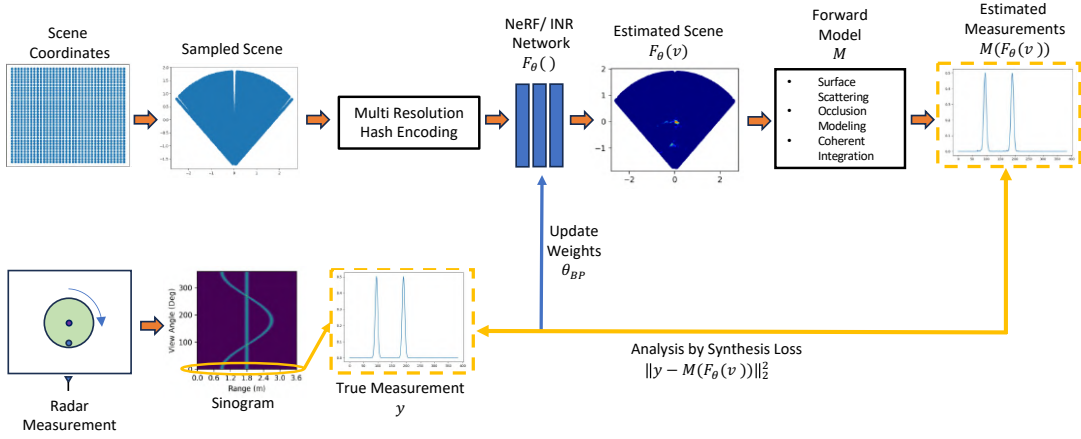


Fig. 2: illustrates our proposed analysis-through-synthesis (ATS) pipeline for radar image reconstruction. First, scene coordinates are sampled using our spherical sampling scheme and encoded with a multi-resolution hash encoding. Next, the encoded coordinates are fed into the NeRF network to predict scene scattering functions, which are then used to generate radar measurements through a differentiable forward model. Training the network involves minimizing the loss between estimated and true radar measurements for each scan of the sinogram.

an electronically motorized turn-table for placing the target object and getting a steady rotational motion. Additionally, we demonstrate the feasibility of collecting data with a standard floor cleaning robot by integrating supplementary circuitry (costing approximately \$10) [22], which incorporates an ESP32 chip with Bluetooth functionality.

In this setting, radar measurements are acquired from n virtual radar positions, or in our case, from 360° positions creating a circular synthetic aperture. These radar measurements are aggregated to create a 2D sinogram, where the x-axis represents the range in meters, and the y-axis corresponds to ϕ in degrees. Each ϕ value aligns with a radar position or a 1° rotation angle of the turn-table. While the measurements are intended to be static for each angle, in practice, data is continuously collected until the rotating table completes a full revolution. The only requirements are to maintain a constant rotational speed of the target during the sensing period and to keep the axis of rotation unchanged throughout the observation period.

IV. ANALYSIS-THROUGH-SYNTHESIS (ATS) PIPELINE

After acquiring the measurements (i.e. sinograms), our reconstruction process employs an analysis-through-synthesis optimization technique, leveraging an implicit neural representation (INR) similar to NeRF, in traditional view synthesis. Figure 2 shows our proposed analysis-through-synthesis (ATS) pipeline. At first, the scene coordinates are sampled using our spherical sampling scheme for a given sensor position. Then the sampled scene is passed through a multi-resolution hash encoding block similar to Instant-NGP [31] for generating the positional embedding v from each coordinate. It uses hash functions to generate varied positional embedding, capturing relative positions in a sequence at different resolutions. In contrast to positional encoding, the multi-resolution approach allows a more dynamic and flexible

representation of positional information. Then the encoded coordinates are fed into the NeRF neural network of four fully connected layers, referred to as the neural backprojection network \mathcal{N}_{BP} , parameterized by weights θ_{BP} . Through this network, we predict the complex scattering function σ' in the estimated scene $F_\theta(v)$. More precisely, the network characterizes the scatterers at each location of the scene. The scatterers within the scene are subsequently employed to generate radar measurements (each scan in the sinogram) through our differentiable forward model described in section IV-A. Training the network involves minimizing the loss between the synthesized measurements $M(F_\theta(v))$ and actual radar measurements y for each scan of the sinogram. In short, by employing a neural network, we predict scene scatterers, and then by employing a differentiable forward model, we synthesize radar measurements over time.

A. Radar Forward Model

In this subsection, we present the radar forward measurement model, which serves as the foundation for our ATS pipeline. This model draws inspiration from a point-based radar scattering model as outlined in [32], [33]. Such point-scattering models provide computational feasibility and differentiability, making them conducive to neural rendering.

Let $x \in \mathbb{R}^3$ describe a 3D coordinate in a scene, $\sigma(x) \in \mathbb{R}$ denote the amplitude of the scatterer at x in the scene, and $s(t)$ represent the transmitted UWB signal:

$$s(t) = e^{-\frac{t^2}{2\tau_0^2}} \cdot \cos(2\pi f_c t) \quad (1)$$

Here, f_c is the center frequency of the transmitted pulse, and τ_0 is the standard deviation that controls the pulse width of the UWB signal.

Figure 3 illustrates the forward model geometry, with transmitter and receiver origins designated as o_T and o_R

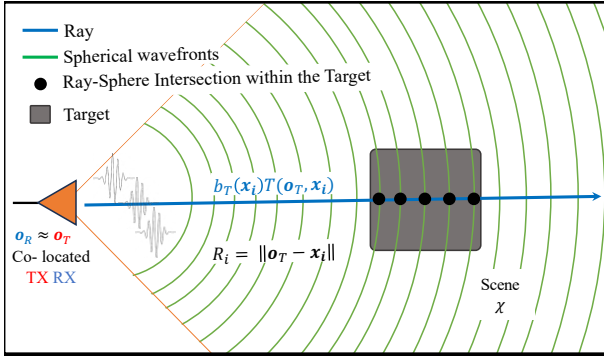


Fig. 3: Illustrates our forward model geometry and sampling strategy. A transmitted ray (blue) is emitted towards the scene, propagating to a scene χ weighted by the directivity function $b_T(x)$ and transmission probability $T(o_T, x)$. All transmitted rays within the antenna beamwidth (orange lines) are then sampled at the intersection point of the sphere (green) defined by range samples. TX and RX are collocated.

respectively. $b_T(x)$ represents the directivity function for the transmitter and $T(o_T, x)$ is defined as the transmission probability between a point in the scene and the origins of the transmitter and receiver respectively, accounting for occlusion effects in our model.

Let $R = \|o_T - x\| = \|o_R - x\|$ denote the distance between the point in the scene and the origins of the transmitter and receiver. Then the receiver signals can be expressed as:

$$r(t) = \int_{\chi} \frac{2T(o_T, x)b_T(x)}{(4\pi)^3 R^4} L(\sigma(x)) s\left(t - \frac{2R}{c}\right) dx \quad (2)$$

where χ denotes the set of all coordinates in the region of interest in the scene, $L(\sigma(x))$ represents the Lambertian scattering model of the scene point, and c is the speed of light. Considering the co-location of transmitter and receiver antennas, we assume equal transmission and reception probabilities, hence the factor of 2 in the numerator.

However, pulses are transmitted and received within discrete radial range bins in ultra-wideband radar systems. These distances are often referred to as sets of points lying on constant time-of-flight paths, also known as radial wavefronts. Received pulses are coherently integrated based on their constant time-of-flight range bins, thereby forming radar range profiles. Consequently, we approximate the forward model similar to [11] that captures these characteristics given by the following equation:

$$r\left(t = \frac{2R}{c}\right) = \int_{E_r} 2T(o_T, x)b_T(x)L(\sigma(x))dx. \quad (3)$$

Here, E_r delineates the sphere encompassing all points within the constant time of flight for the radar's forward model geometry, $\sigma(x)$ denotes the scene scatterer at x , and its complex scattering function is estimated by a neural network, represented as $\sigma(x) = \mathcal{N}_{BP}(x; \theta_{BP})$.

B. Spherical Sampling

In this subsection, we outline our method of sampling the scene using spheres of constant time-of-flight to estimate the transmission probabilities for transmitted rays.

In Figure 3 each green semi-circle corresponds to a constant time-of-flight $t = 2R/c$ which forms a sphere. This sphere is centered where the transmitter and receiver antennas coincide ($O_T \simeq O_R$), with a radius of $R = ct/2$. The equation representing this sphere is:

$$x^2 + y^2 + z^2 = R^2. \quad (4)$$

Now, by intersecting a sphere with radius $\frac{ct}{2}$ with a bundle of rays originating from the transmitter and falling within the beamwidth θ_{BW} , we obtain the sampled points in the scene. The transmission ray, depicted in blue with direction \vec{d}_{T_j} , is defined as

$$x_{T_{i,j}} = o_T + l_i \cdot \vec{d}_{T_j} \quad (5)$$

where l_i represents the depth samples along the ray calculated at which a ray intersects the sphere by substituting the ray into Equation 4. This substitution results in a quadratic equation from which we extract the positive root:

$$x = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad (6)$$

where,

$$a = \vec{d}_T \cdot \vec{d}_T, b = 2\vec{x}_T \cdot \vec{d}_T, c = \vec{x}_T \cdot \vec{x}_T - R^2 \quad (7)$$

We address occlusion by calculating transmission probabilities between the transmitter/receiver and various scene points using the method of Mildenhall et al. [8] and Reed et al. [11]. These probabilities are expressed as the product of exponential terms, where $|\sigma_k|$ represents the magnitude of the scattering coefficient at range bin k , and $|l_{k+1} - l_k|$ denotes the distance between consecutive range bins:

$$T(o, x_{T_i}) = \prod_{k < i} e^{-(|\sigma_k| \cdot |l_{k+1} - l_k|)}.$$

Utilizing the Lambertian scattering model, the scattered intensity $L(\sigma)$ is calculated as:

$$L(\sigma) = \sigma \cdot \frac{(x_{T_i} - o_T)}{\|x_{T_i} - o_T\|} \cdot 2T(o, x_{T_i}).$$

These equations facilitate the computation of transmission probabilities and scattered intensity, crucial for addressing occlusion in the scene.

V. SYNTHETIC DATA EXPERIMENTS

To generate synthetic radar sinograms depicting various point targets, we employ Equation 1. Four distinct synthetic scenarios were simulated, each containing single, double, triple, and quadruple targets within the scene. The corresponding results are presented in Figure 4. ATS demonstrates qualitative consistency with BP, exhibiting accurate scene reconstruction in noise-free conditions and superior performance in noisy conditions. BP, on the other hand, shows strong artifacts around the corners of each reconstructed scene when noise is added.

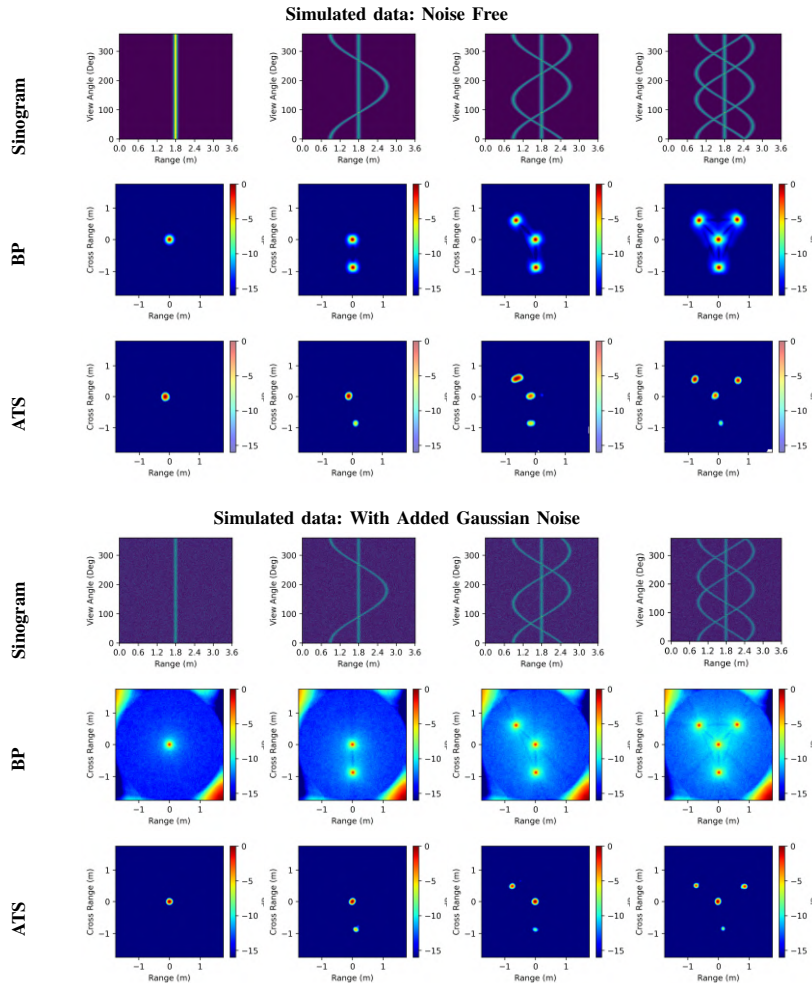


Fig. 4: Simulated data in noise-free and noisy conditions. The rows show the Sinogram, BP, and ATS reconstructions respectively for single, double, triple, and quad point targets.

TABLE I: Quantitative Metrics Comparison between different Reconstruction Methods on Simulated Data with added Gaussian Noise $CN \sim \mathcal{N}(0, 0.1)$

# Reflectors	Method	PSNR \uparrow	LPIPS \downarrow	MSE \downarrow
One	BP	11.797	1.055	0.062
	ATS	15.049	0.069	0.001
Two	BP	11.296	1.050	0.069
	ATS	14.611	0.047	0.004
Three	BP	10.754	1.043	0.077
	ATS	14.134	0.069	0.004
Four	BP	10.266	1.035	0.084
	ATS	13.746	0.118	0.004

A. Effect of Noise

To quantitatively evaluate our proposed ATS reconstruction against BP, we use three image metrics namely- PSNR (Peak Signal to Noise Ratio), LPIPS (Learned Perceptual Image Patch Similarity) [34], and MSE (Mean Squared Error). Table I presents the comparison of quantitative metrics among various reconstruction methods on four different simulated scenes with added Gaussian noise $CN \sim \mathcal{N}(0, 0.1)$ per scan. As shown in Table I, the proposed ATS consistently

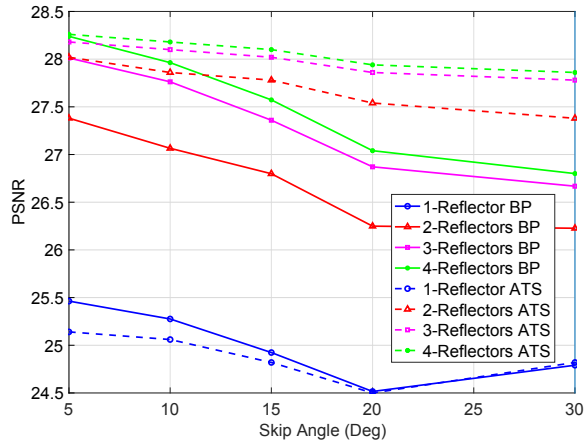
outperforms BP across all metrics in noisy scenes.

B. Effect of Sparse Measurements

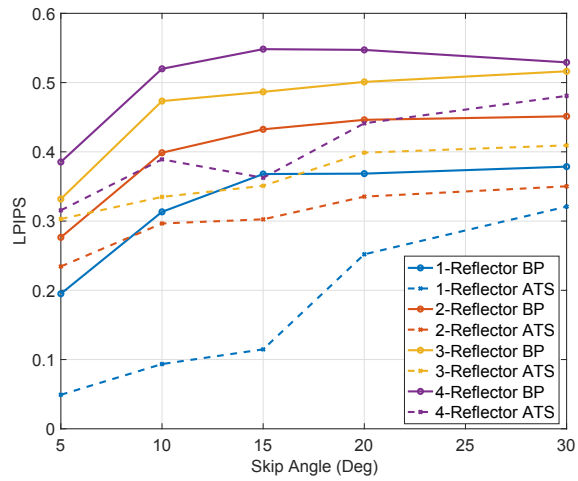
Figures 5a and 5b illustrate the influence of sparse measurements in terms of skip angles on PSNR and LPIPS metrics, respectively, for both BP and ATS across varying numbers of simulated targets. With an increasing number of targets, PSNR tends to decrease for both BP and ATS; however, ATS consistently maintains higher PSNR values compared to BP across all skip angles except for the single reflector case where the PSNR difference is less than 0.4. Similarly, ATS exhibits lower LPIPS values compared to BP across all skip angles, indicating superior performance in scenarios with limited measurements.

VI. REAL DATA EXPERIMENTS

Figure 6 illustrates ISAR imaging results obtained with real data collected with - a single cylindrical object, two cylindrical objects placed at varying distances apart, and a square container. The second and third columns of the figure illustrate reconstructed images using BP and ATS, respectively. BP begins to exhibit artifacts as the complexity



(a) PSNR \uparrow comparison for BP and ATS



(b) LPIPS \downarrow comparison for BP and ATS

Fig. 5: Comparison of PSNR and LPIPS metrics for BP and ATS with different numbers of simulated reflectors at various skip angles

of the target scene increases. In contrast, ATS reconstruction displays no artifacts and maintains the reconstructed targets' dimensions more accurately to their original form.

A. Effect of Skip Angles on Real Scene Reconstruction

The impact of measurements with skip angles on the reconstruction algorithms for imaging double soda cans is depicted in Figure 7. Increasing skip angles introduces significant artifacts in BP reconstruction, whereas ATS reconstruction only suffers from minor amplitude attenuation.

B. Effect of Partial Rotations on Real Scene Reconstruction

The effect of partial rotation of the target on BP and ATS reconstructions of double soda cans is shown in Figure 8. It can be observed that BP fails to reconstruct the secondary object when the measurements are less than or equal to half of the total 360 measurement angles.

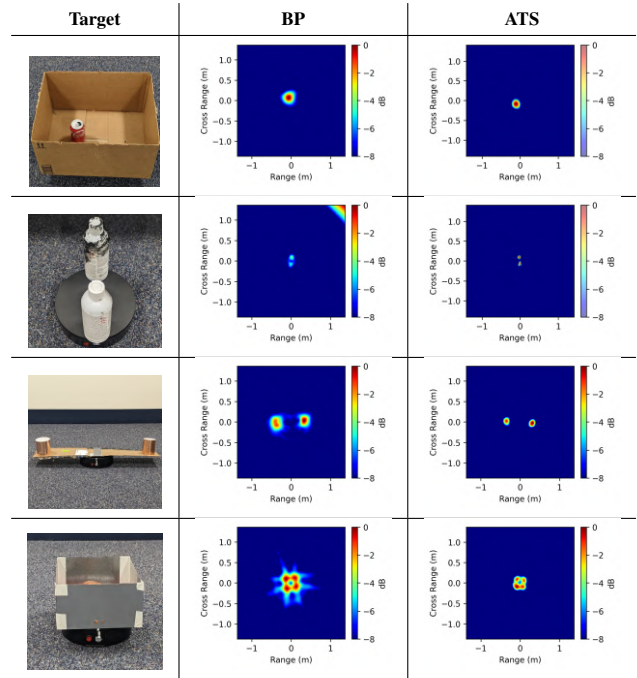


Fig. 6: Real data (with full 360° views with 1 scan per viewing angle) - Row 1: single soda can inside a cardboard box, Row 2: two spray paint bottles 7 inches (17.78 cm) apart, Row 3: double cans 27 inches (68.58 cm) apart, Row 4: square metal box with a side length of 9.5 inches (24.13 cm).

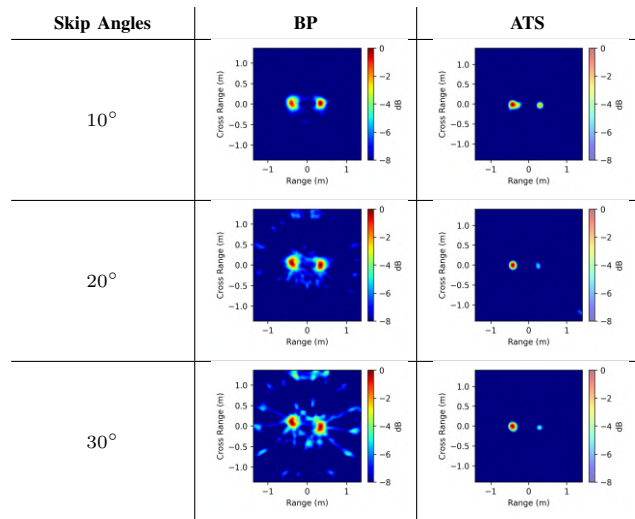


Fig. 7: Impact of skip angles on double soda cans imaging. Row 1: 10° skip (36 scans), Row 2: 20° skip (18 scans), Row 3: 30° skip (12 scans). Increasing skip angles leads to notable artifacts in BP. In contrast, ATS reconstruction experiences only a slight decrease in signal strength.

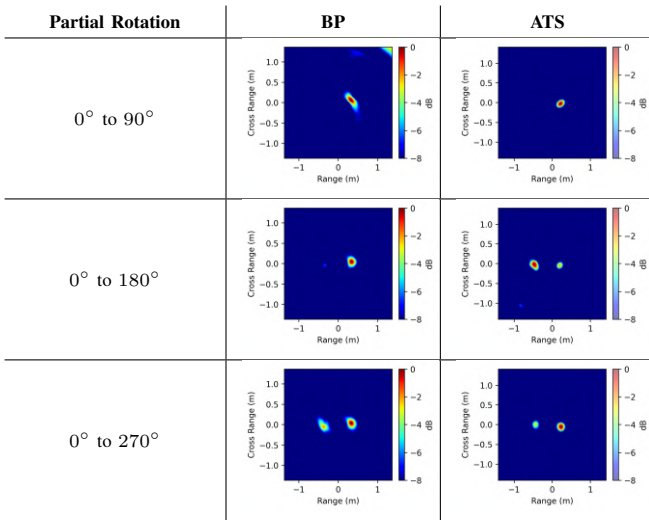


Fig. 8: Effect of limited measurement angles on double soda cans imaging. Row 1: 90 scans (from 0° to 90°), Row 2: 180 scans (from 0° to 180°), and Row 3: 270 scans (from 0° to 270°). BP fails to reconstruct the secondary object when measurements are $\leq 180^\circ$.

VII. APPLICATIONS OF ATS

A. Identification Applications:

Numerous studies have investigated embedding information into objects for non-intrusive content recognition. Radar-based methods like MechanoBeat [35], for instance, utilize the oscillation frequency of 3D-printed tags to embed identification information into everyday objects. Motivated by this concept, our objective in the subsequent experiments was to integrate identification information into objects such as cylindrical containers and helmets by equipping them with 3D-printed corner reflectors. We maintained a consistent number of reflectors to ensure a constant Radar Cross-Section (RCS) while varying their orientations to enhance robust detection capabilities. We utilize 3D-printed octahedral corner reflectors (20 mm radius) positioned in specific orientations to encode identification information within objects. We apply copper tapes onto the surfaces of the reflectors to further enhance their reflectivity.

1) **Plastic Container/Mug Identification:** The top two rows of Figure 9 displays imaging of a 3D-printed plastic cylindrical container having two corner reflectors positioned at 90° and 180° orientations respectively. The results show that the ATS method can reasonably reconstruct the relative positions of reflectors. In contrast, BP presents noticeable artifacts, failing to provide clear tag reconstruction even after a -30dB dynamic range adjustment.

2) **Helmet Identification:** In challenging environments such as construction sites and firefighting scenarios where line-of-sight may be obstructed, accurate ISAR imaging facilitates easy identification of helmets through the attached tags. Figure 9 (bottom two rows) presents imaging results of helmets fitted with 3D-printed reflector tags at varying orientations, showcasing the effectiveness of the ATS method

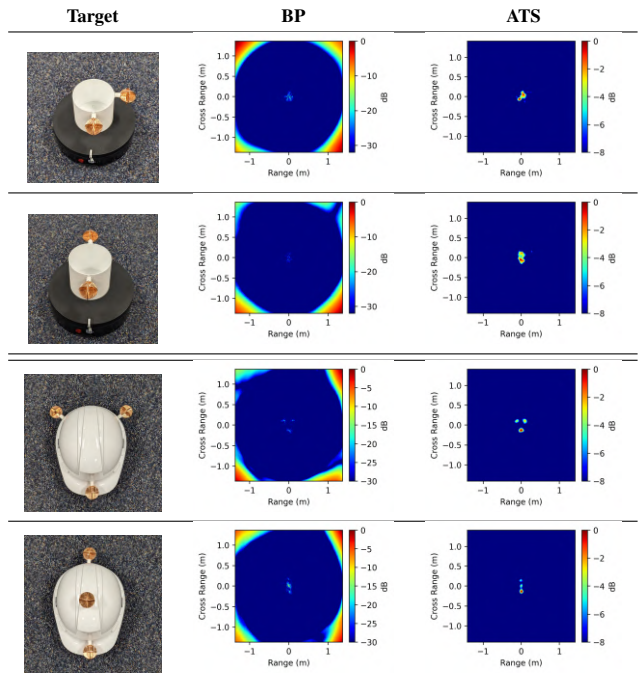


Fig. 9: Enhanced object identification using 3D printed octahedral reflectors: Top two rows depict ISAR imaging of a cylindrical container with corner reflectors placed at 90° and 180° orientations, while bottom two rows showcase ISAR imaging of a construction helmet with reflectors positioned at 120° intervals and in a linear alignment.

in accurately reconstructing reflectors' positions. In contrast, BP exhibits notable artifacts and inferior tag imaging, visible only after a -30dB dynamic range adjustment.

B. Potential Imaging Applications:

The potential applications of ATS with radar technology span various fields. For instance, in robot-environment interaction, ATS enables robots equipped with SAR capabilities to thoroughly explore objects through scanning and reconstruction, thereby enhancing environmental understanding. In security scenarios, ATS can penetrate concealments to reconstruct the internal structures of complex objects, aiding in security clearance procedures by detecting concealed weapons. Additionally, ATS could be used for unlabelled warehouse package recognition and reconstruction, allowing for the identification of package contents without repackaging.

VIII. CONCLUSION AND FUTURE WORK

In conclusion, we have presented an ISAR imaging algorithm that adopts an analysis-through-synthesis approach inspired by conventional NeRF techniques. By estimating scene scatterers through an implicit neural network and optimizing with a differentiable forward model incorporating radar wave propagation and reflection characteristics, our analysis-through-synthesis framework demonstrates quantitative and qualitative superiority over traditional backprojection through extensive experiments on simulated and

hardware-measured data. Moreover, our approach eliminates the need for costly anechoic chambers or complex measurement testbeds, reducing both cost and computation time in the reconstruction process. Leveraging the penetration capability of our radar forward model, we achieve reconstruction in Non-Line-of-Sight scenarios such as imaging contents inside a cardboard box.

Although ATS demonstrates superior performance over BP in scenarios with noise and limited measurement space, its drawback lies in its lack of real-time capability. Particularly in real-time applications where latency and speed are crucial, ATS is still considered slow (4.36 seconds) compared to BP (0.29 seconds) using NVIDIA RTX 3080Ti GPU. Exploring strategies such as sub-sampling or skipping may offer potential acceleration in measurement capture, warranting further investigation. Additionally, extending the framework to enable 3D volumetric reconstruction could open up possibilities for imaging applications such as robot environment interaction, security scanning, and warehouse package imaging.

REFERENCES

- [1] C. J. Li and H. Ling, "An investigation on the radar signatures of small consumer drones," *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 649–652, 2017.
- [2] C. J. Li and H. Ling, "Wide-angle, ultra-wideband isar imaging of vehicles and drones," p. 3311, 2018.
- [3] A. R. Brenner and L. Roessing, "Radar imaging of urban areas by means of very high-resolution sar and interferometric sar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 10, pp. 2971–2982, 2008.
- [4] C. Zhao, L. Xu, X. Bai, and J. Chen, "Near-field high-resolution sar imaging with sparse sampling interval," *Sensors*, vol. 22, no. 15, 2022.
- [5] M. E. Yanik, D. Wang, and M. Torlak, "Development and demonstration of mimo-sar mmwave imaging testbeds," *IEEE Access*, vol. 8, pp. 126 019–126 038, 2020.
- [6] T. Sakamoto, T. Sato, P. J. Aubry, and A. G. Yaroyov, "Ultra-wideband radar imaging using a hybrid of kirchhoff migration and stolt f-k migration with an inverse boundary scattering transform," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 8, pp. 3502–3512, 2015.
- [7] L. Kang, Y. Luo, Q. Zhang, X.-W. Liu, and B.-S. Liang, "3-d scattering image sparse reconstruction via radar network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.
- [8] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.
- [9] R. Martin-Brualla, N. Radwan, M. S. Sajjadi, J. T. Barron, A. Dosovitskiy, and D. Duckworth, "Nerf in the wild: Neural radiance fields for unconstrained photo collections," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 7210–7219.
- [10] M. Niemeyer, J. Barron, B. Mildenhall, M. S. M. Sajjadi, A. Geiger, and N. Radwan, "RegNeRF: Regularizing neural radiance fields for view synthesis from sparse inputs," in *Computer Vision and Pattern Recognition (CVPR)*, 2022.
- [11] A. Reed, J. Kim, T. Blanford, A. Pediredla, D. Brown, and S. Jayasuriya, "Neural volumetric reconstruction for coherent synthetic aperture sonar," *ACM Transactions on Graphics (TOG)*, vol. 42, no. 4, pp. 1–20, 2023.
- [12] A. Reed, T. Blanford, D. C. Brown, and S. Jayasuriya, "Implicit neural representations for deconvolving sas images," in *OCEANS 2021: San Diego-Porto*. IEEE, 2021, pp. 1–7.
- [13] A. Reed, T. Blanford, D. C. Brown, and S. Jayasuriya, "Sinr: Deconvolving circular sas images using implicit neural representations," *IEEE Journal of Selected Topics in Signal Processing*, vol. 17, no. 2, pp. 458–472, 2023.
- [14] S. M. Farrell, V. Boominathan, N. Raymondi, A. Sabharwal, and A. Veeraraghavan, "Coir: Compressive implicit radar," *IEEE Trans on Pattern Analysis and Machine Intelligence*, pp. 1–12, 2023.
- [15] A. Liu, S. Zhang, C. Zhang, S. Zhi, and X. Li, "Ranerf: Neural 3-d reconstruction of space targets from isar image sequences," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–15, 2023.
- [16] J. Deng, P. Xie, L. Zhang, and Y. Cong, "Isar-nerf: Neural radiance fields for 3d imaging of space target from multi-view isar images," *IEEE Sensors Journal*, pp. 1–1, 2024.
- [17] H. Zhang, Y. Lin, F. Teng, S. Feng, B. Yang, and W. Hong, "Circular SAR incoherent 3d imaging with a nerf-inspired method," *Remote Sensing*, vol. 15, no. 13, 2023.
- [18] S. Zhou, S. Xie, R. Ishikawa, K. Sakurada, M. Onishi, and T. Oishi, "Inf: Implicit neural fusion for lidar and camera," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023, pp. 10918–10925.
- [19] Q. Herau, N. Piasco, M. Bennehar, L. Roldão, D. Tsishkou, C. Migniot, P. Vasseur, and C. Demonceaux, "Moisst: Multimodal optimization of implicit scene for spatiotemporal calibration," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023, pp. 1810–1817.
- [20] X. Zhong, Y. Pan, J. Behley, and C. Stachniss, "Shine-mapping: Large-scale 3d mapping using sparse hierarchical implicit neural representations," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 8371–8377.
- [21] JAYEGT, "Motorized rotating display stand," 2022. [Online]. Available: <https://www.amazon.com/JAYEGT-Motorized-Rotating-Turntable-Photography/dp/B08K35HPJ4>
- [22] "Roomba more smarter," 2024, <https://kidikeita.com/posts/20240308-roomba/>.
- [23] O. Frey, E. Meier, and D. Nüesch, "A study on integrated sar processing and geocoding by means of time-domain backprojection," in *Proceedings of the Int. Radar Symposium, Berlin*. Citeseer, 2005.
- [24] J. S. S. G. T. Benjamin and C. Flores, *Range-Doppler Radar Imaging and Motion Compensation*. Artech House, 2001.
- [25] D. Zhu, S. Ye, and Z. Zhu, "Polar format algorithm using chirp scaling for spotlight sar image formation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 4, pp. 1433–1448, 2008.
- [26] C. Hu, L. Wang, Z. Li, and D. Zhu, "Inverse synthetic aperture radar imaging using a fully convolutional neural network," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 7, pp. 1203–1207, 2019.
- [27] S. Tan, X. Zhang, H. Wang, L. Yu, Y. Du, J. Yin, and B. Wu, "A cnn-based self-supervised synthetic aperture radar image denoising approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2021.
- [28] W. Li, Y. Yuan, Y. Zhang, and Y. Luo, "Unblurring isar imaging for maneuvering target based on ufgan," *Remote Sensing*, vol. 14, no. 20, p. 5270, 2022.
- [29] L. Wang, M. Zheng, W. Du, M. Wei, and L. Li, "Super-resolution sar image reconstruction via generative adversarial network," in *2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE)*. IEEE, 2018, pp. 1–4.
- [30] "Pulson 440 uwb radar," 2015, <https://fccid.io/NUF-P440-A/User-Manual/User-Manual-287844>.
- [31] T. Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," *ACM Transactions on Graphics (ToG)*, vol. 41, no. 4, pp. 1–15, 2022.
- [32] B. Poole, A. Jain, J. T. Barron, and B. Mildenhall, "Dreamfusion: Text-to-3d using 2d diffusion," *arXiv preprint arXiv:2209.14988*, 2022.
- [33] Z. Chance, A. Kern, A. Burch, and J. Goodwin, "Differentiable point scattering models for efficient radar target characterization," *arXiv preprint arXiv:2206.02075*, 2022.
- [34] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 586–595.
- [35] M. F. Tasnim Oshim, J. Killingback, D. Follette, H. Peng, and T. Rahman, "Mechanobeat: Monitoring interactions with everyday objects using 3d printed harmonic oscillators and ultra-wideband radar," in *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, ser. *UIST '20*. New York, NY, USA: Association for Computing Machinery, 2020, p. 430–444. [Online]. Available: <https://doi.org/10.1145/3379337.3415902>