Classification Of Automotive Targets Using Inverse Synthetic Aperture Radar Images

Neeraj Pandey, *Student Member, IEEE*, Shobha Sundar Ram, *Senior Member, IEEE*

Abstract—Traditionally, point cloud representations or Doppler spectrograms have been generated from short-range automotive radars for dynamic object detection and classification. In this work, we propose using inverse synthetic aperture radar (ISAR) images obtained from range compensated turning targets for the classification of different types of vehicles. We experimentally demonstrate that ISAR images of automotive targets provide rich features such as the dimensions, trajectory, and the number of wheels of the vehicles for classification. Additionally, we present a simulation framework for generating large volumes of realistic ISAR images of automotive targets at millimeter-wave frequencies for training classifiers. The model incorporates radar scattering phenomenology of commonly found vehicles along with range-Doppler-based clutter and receiver noise. The model is experimentally validated with measurement data gathered from an automotive radar. The images from the simulation database are subsequently classified using traditional machine learning techniques and deep neural networks based on transfer learning. We show that the ISAR images offer a classification accuracy above 90% and are robust to both noise and clutter.

Index Terms—ISAR, classification, automotive radar, transfer learning, radar database

I. INTRODUCTION

With the advent of advanced driver assistance systems (ADAS), automotive radars are becoming increasingly common on cars for improving road driving conditions. These radars are used for multiple applications such as automatic cruise control, pedestrian detection, cross-traffic alert, blind-spot detection, and parking assistance [1], [2]. The main advantage of automotive radar over the camera for object detection and classification is that the radar can be operated in low light conditions, rain, and fog. Secondly, these sensors are typically cheaper than cameras, and hence multiple of them can be mounted around the periphery of the vehicle, usually behind the bumpers. Finally, automotive radars operate at millimeter-wave frequencies with high bandwidths and spatially large antenna arrays. Hence, they offer an excellent range, Doppler, and azimuth resolution [3]. Usually, in these systems, the raw radar data cube is processed to provide a collection of point scatterers corresponding to both vehicles and clutter with range, azimuth, elevation, and Doppler information. Direct object detection and classification based on this type of data can be challenging since it is difficult to correctly cluster the point scatterers are belonging to the same object [4], [5]. Instead, radar images / signatures directly processed from raw radar data provide more effective features for automatic target recognition.

Classification of radar targets for a variety of applications has been researched over the last few decades [6]–[12]. Many different types of radar signatures have been studied. For example, radar micro-Doppler spectrograms, which are the joint time-frequency representations of time-domain narrowband radar data have proven to provide excellent features for classification. They have been used for distinguishing between different types of human activities [13]–[18]; armed and unarmed personnel [19]; ground vehicles and pedestrians [20]–[22]; and different types of airborne targets such as unmanned

N.P and S.S.R. are with the Indraprastha Institute of Information Technology Delhi, New Delhi 110020 India. E-mail: {neerajp, shobha}@iiitd.ac.in.

aerial vehicles and birds [23], [24]. Range-Doppler plots generated from broadband radar data have also served as excellent features for target recognition [25]. Other works have used range-crossrange images generated through synthetic aperture radar (SAR) imaging for classification purposes [26]–[28]. However, SAR images are typically more suited for classifying static targets since dynamic targets may distort the radar images. The alternative is to use antenna array processing for obtaining fine cross-range resolution. However, this requires a large array with lots of antenna elements, and precise phase synchronization across the multiple channel data [29]. A third method for obtaining fine cross-range resolution is to use inverse synthetic aperture radar (ISAR) processing of single-channel broadband data.

When a dynamic target travels along a complex trajectory, the target undergoes a combination of translational and rotational motion. If the translational motion of a target can be correctly estimated and compensated, then the Doppler dimension can be mapped to cross-range to obtain ISAR images [30]. The cross-range resolution is inversely proportional to the extent of the target aspect presented to the radar during the rotational motion and the coherent processing time interval. ISAR images, generated from range-Doppler plots of dynamic targets, have been extensively researched over the last two decades - especially for the detection and classification of airborne targets and ships [31]–[34]. In the automotive target scenario, vehicles undergo complex turns, which can result in a large radar aspect. Further, even while moving along a straight path, a slight offset of the target vehicular trajectory from the ego radar, can result in large radar aspects to get a fine cross-range resolution. Since automotive short-range radars are characterized by large bandwidths (above 2GHz) and high carrier frequencies (77GHz) that result in fine range and Doppler (or cross-range) resolution, they are particularly suited for generating high-resolution ISAR images of vehicles. In ADAS systems, multiple auxiliary sensors (gyrometers, accelerometers, other radars) are deployed on the ego vehicle. Therefore, the translational motion compensation of both the ego vehicle and target vehicle can usually be carried out without too much difficulty. More recently, ISAR images of ground-based targets have been generated using turntable data [35], ground-based platforms [36], [37], and from airborne platforms [38]. However, these studies have been restricted to very few targets. In our preliminary paper in [39], we showcased how these images provide detailed insights into the dimensions of vehicles and their trajectories. However, the images were idealized and free of corruption from noise and ground clutter.

Our contributions in this paper are as follows:

1) We provide a detailed simulation framework to generate realistic ISAR images of automotive targets while incorporating the effects of additive noise in the radar receiver and speckle noise due to ground clutter effects. The main objective is to provide a simulation framework for rapidly generating large volumes of radar data without the cost and man-hours involved in collecting measurement data. These data can be used for training deep neural networks, which have recently emerged as the algorithm of choice for classifying radar images [14], [17], [27]. Secondly, the simulation framework

can be integrated with software test beds for rapid prototype development and validation. Finally, the simulation models can be useful for understanding radar propagation phenomenology in the environment and pinpointing cause and effect. We have considered five commonly found targets - a full-size car, a mid-size car, a four-wheel truck, a bicycle, and an auto-rickshaw (tuk-tuk). We generate ISAR radar images of these targets performing different types of turns (right, left, and U-turn) as well as following a straight trajectory.

2) In real-world conditions, there can be considerable clutter arising from the rough road surface at millimeter-wave frequencies, which is proportional to the radar coverage area [40]. Hence, this range-based clutter cross-section increases with radar range. Further, Doppler-based clutter can arise due to wind [41]. The combination of range and Doppler-based clutter manifests as speckle noise and can significantly distort the ISAR images. In this work, we have incorporated detailed range-Doppler clutter as well as receiver noise in our radar models to simulate realistic radar images. With this paper, we release our database, consisting of over 30000 realistic ISAR images, to the radar community at https://tinyurl.com/msu6aj7y.

3) Third, we demonstrate that these images show detailed information of the type of vehicle, its dimensions, the number of wheels, and the trajectory followed by the vehicle. Further, we have validated these simulated images with measurement data gathered from Texas Instrument's AWR 1843 77GHz automotive radar.

4) Finally, most importantly, we demonstrate that the ISAR images offer distinctive features for the classification of automotive targets - compared to traditional features like point cloud representations and Doppler spectrograms. To showcase the effectiveness of ISAR features, we have considered both traditional machine learning algorithms such as support vector machine (SVM) [42] and random forest (RF) [43] as well as Alexnet and Googlenet, which are two transfer learning based deep neural networks [44]. Our results show that the ISAR images are successfully classified by the machine learning algorithms (with a precision and recall above 90%). The deep neural networks outperform the traditional machine learning algorithms and are robust to noise and clutter.

The paper is organized as follows. In the Section II , we present the simulation methodology for modeling the scattered signal radar signals from the automotive targets, as well as the noise and clutter models. Then we describe the radar signal processing algorithms for generating the ISAR images. In Section III, we present the experimental set up for collecting measurement data for generating ISAR images and present the measurement results. In Section IV, we present the classification results of the five automotive targets using four different machine learning-based algorithms - SVM, RF, Alexnet, and Googlenet. Finally, we conclude the paper with our final remarks in Section V.

II. SIMULATION METHODOLOGY

While several prior works have described simulation models of pedestrians [45], [46], there are very few works that model automotive vehicles along complex trajectories [47], [48]. These works have confined their scope to simulating high range resolution profiles and micro-Doppler spectrograms. In this section, we discuss the simulation methodology for modeling the scattered radar signals from five common automotive targets - bicycle, auto-rickshaw, mid-size car, full-sized car, and truck. Then we describe how these models

can be integrated into the radar waveform to obtain ISAR images. Finally, we present the method to incorporate noise and clutter in the images.

A. Automotive Animation Model

We imported freely available three-dimensional (3D) computer aided design (CAD) models[1] of the automotive targets into Blender software. Then, we rendered the metallic parts of the automotive into triangular facets. An accurate rendering of the target capturing the diversity of features on the chassis of the vehicle is realized by using a large number of facets. In our work, the bicycle, and auto-rickshaw are rendered with 3919 and 6949 facets, respectively; the mid-size and full-size cars with 6905 and 19964 facets, respectively, while the truck is rendered with 7206 facets as shown in Fig.1. We have
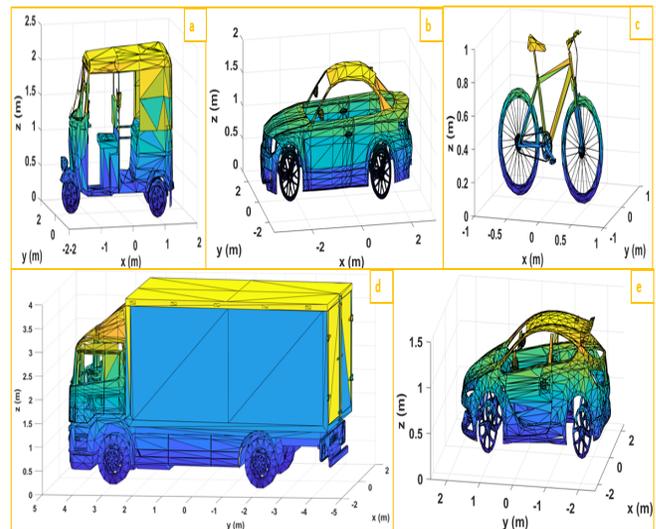
Fig. 1: Three-dimensional automotive targets with triangular facets used for the simulation. (a) Auto-rickshaw (tuk-tuk), (b) full-sized car (c) bicycle, (d) truck (e) mid-sized car

considered a four-way traffic junction, where lanes from the north (N), south (S), east (E), and west (W) meet as shown in Fig2a. The
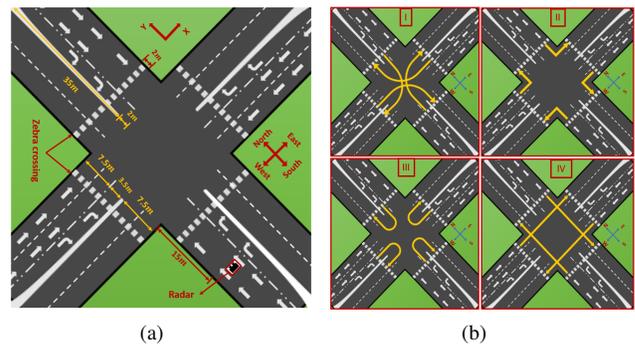
Fig. 2: (a) Road geometry of four-way traffic junction. (b) Trajectories undertaken by the automotive target in a four-way junction - (i) Right turn, (ii) Left turn, (iii) U-turn and (iv) Straight through.

targets are assumed to stand on the XY ground plane which is aligned with the N-S and E-W directions with the height along the Z-axis. The ego radar is assumed to be static and fixed at $(0, 0, 0.5)$m along the south road. A total of 16 different trajectories are possible at this junction. They are the four right turns (S to E, E to N, N to W, W

[1] https://free3d.com/3d-models/

to S), four left turns (S to W, W to N, N to E, E to S), four U-turns (S to S, E to E, N to N, W to W) and four straight through (S to N, N to S, W to E, E to W) as shown in Fig2b.

Next we describe the method for animating the vehicle along the desired trajectory at a specified speed. We first identify the center of the vehicle $\vec{r}_C$ and fix it at the starting position along a trajectory. Then, we identify way points along the distance of the trajectory such that the time taken for the vehicle's center to travel between any two way points is fixed (t_f) as shown in Fig.3a. The sampling
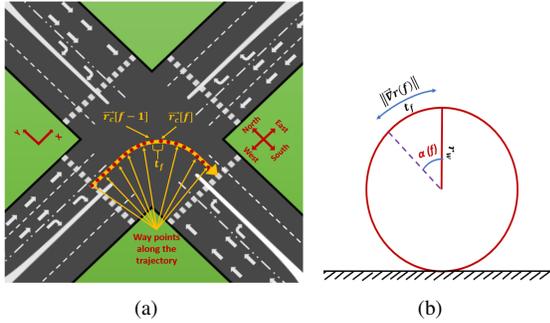
Fig. 3: Animation model of trajectory: (a) Way points along the trajectory that will be traversed by the center of the vehicle; (b) wheels rotational angle calculation.

time instants corresponding to these way points are indicated by $f = 1, 2, \cdots, F$. Therefore, the center of the vehicle undergoes translational displacement $\vec{\nabla}r[f] = \vec{r}_C[f] - \vec{r}_C[f - 1]$ at each f. The target is composed of B triangular facets where the centroid of each facet is $\vec{r}_b, b = 1 : B$. Based on the rendering of the vehicle, we obtain displacement vectors ($\vec{\nabla}r_b$) of these centroids from the center of the vehicle. These displacement vectors are fixed with time since the chassis of the vehicle behaves as a rigid body (excepting the wheels). The vehicle must undergo rotational motion along with translational motion. This rotational motion at each frame f is described in terms of the yaw (rotation angle about the Z axis, $\theta_b[f]$) which is computed from

$$\theta[f] = \arctan\left(\frac{y_C[f] - y_C[f - 1]}{x_C[f] - x_C[f - 1]}\right). \tag{1}$$

where y_C and x_C are y and x coordinates of centriod.

The position of the facet centroid on the chassis of the vehicle is

$$\vec{r}_b[f] = \left(R_{\theta[f]}\vec{r}_b[f - 1]\right) + \vec{\nabla}r[f], \tag{2}$$

where R_θ is the Euler rotation matrix for a yaw of θ. In the case of wheels, a facet centroid on the wheel undergoes additional rotation due to the motion of the wheel. The angular displacement of the wheel is proportional to the distance travelled by the wheel and the radius of the wheel r_w as shown in

$$\alpha[f] = \frac{||\vec{\nabla}r[f]||}{r_w}. \tag{3}$$

The axis of this wheel rotation is obtained by the cross product of the direction of translational displacement and the height axis. The total displacement of a point on the wheel is

$$\vec{r}_b[f] = \left(R_{\alpha[f]}R_{\theta[f]}\vec{r}_b[f - 1]\right) + \vec{\nabla}r[f], \tag{4}$$

where R_α is the Euler rotation matrix corresponding to a pitch angle of α. The entire algorithm describing the animation motion modeling is summarized in Algorithm.1.

B. Advanced Point Scatterer Signal Model

Automotive radars use linear frequency modulated (LFM) waveforms for the detection and tracking of targets. We consider a radar transmitting an LFM waveform,

---

Algorithm 1: Animation model of vehicle along desired trajectory

Data: Fixed displacement vectors corresponding to facet centroids on the chassis and wheels of the vehicle ($\vec{\nabla}r_b, b = 1 : B$)

1 with respect to the center of the vehicle ($\vec{r}_C$). **Data:** Specify way point positions for center of vehicle along trajectory: $\vec{r}_C[f], f = 1 : F$. Time duration between two way points is fixed (t_f).

Result: Time-varying position coordinates of facet centroids on vehicle ($\vec{r}_b[f], b = 1 : B, f = 1 : F$)

2 Initialization: Initialize positions of all the facet centroids $\vec{r}_b[1] = \vec{r}_C[1] + \vec{\nabla}r_b, b = 1 : B$;

3 **for** $f=2{:}F$ **do**

4   Compute displacement between two consecutive way points $\vec{\nabla}r[f] = \vec{r}_C[f] - \vec{r}_C[f - 1]$;

5   Compute vehicle yaw rotation: $\theta[f] = \frac{y_C[f] - y_C[f - 1]}{x_C[f] - x_C[f - 1]}$;

6   **if** *Facet centroids on chassis* **then**

7   Perform Euler rotation on facet centroids $\vec{r}_b[f] = \left(R_{\theta[f]}\vec{r}_b[f - 1]\right) + \vec{\nabla}r[f], b = 1 : B$;

8   **else if** *Facet centroids on wheels* **then**

9   Compute rotation of wheel $\alpha[f] = \frac{||\vec{\nabla}r[f]||}{r_w}$ where r_w is the radius of the wheel.;

10   Compute axis of wheel rotation which is perpendicular to the plane defined by height axis and vehicle translational motion direction.;

11   Perform Euler rotation on facet centroids on wheel based on wheel rotation axis $\vec{r}_b[f] = \left(R_{\alpha[f]}R_{\theta[f]}\vec{r}_b[f - 1]\right) + \vec{\nabla}r[f], b = 1 : B$

12   **end**

13   **else**

14   **end**

15 **end**

$$s^{tx}(\tau) = rect\left(\frac{\tau}{T_{PRI}}\right)e^{j2\pi f_c\tau}e^{j\pi K\tau^2}, \tag{5}$$

with f_c carrier frequency (and wavelength λ_c) and a chirp rate K. In the above expression, $rect(\cdot)$ indicates that the transmitting signal is defined for a pulse repetition interval of T_{PRI}. We model the automotive target as a collection of moving scattering centers, $b = 1 : B$, located at the centroids of each of the facets on the vehicle. We only model the reflections from the metallic surfaces of the automotive targets since the non-metallic surfaces are known to be poor reflectors [49], [50]. We have strictly considered scenarios where there are single targets in the images. This is usually ensured in real world conditions through pre-processing of raw radar data before classification. In real world scenarios, single channel source separation is applied on multiple targets and returns from each extended target with multiple scattering centers are isolated through range gating and probabilistic association across the temporal history of the data [30], [51], [52]. Since these algorithms are well known to the research community, they are not included in the manuscript due to space constraints. The time-varying range for each scattering center is $r_b = R_b + v_bt$, where R_b is the starting distance from the radar. The radar signal scattered back from a single point target is Doppler shifted by $f_{D_b} = \frac{2v_b}{\lambda_c}$ due to the target's relative radial velocity (v_b) with respect to the radar. The received radar signal, after downconversion to the baseband, is written in terms of slow time t and fast time, τ as

$$s_b^{rx}(\tau, t) = a_b(t)rect\left(\frac{\tau - \tau_b}{T_{PRI}}\right)$$
$$e^{-j\frac{4\pi f_c}{c}R_b}e^{-j2\pi f_{D_b}t}e^{j\pi K(\tau - \tau_b)^2} + \nu, \tag{6}$$

where $\tau_b(t) = \frac{2r_b(t)}{c}$ is the time delay to the target. In the equation above, ν denotes the additive noise that will be discussed in greater detail in the following section. The strength of the received returns, denoted

by a_b, is obtained through the radar range equation by incorporating the transmitted power (P^{tx}), the gains of the transmitting (G^{tx}) and receiving radar antennas (G^{rx}), the radar cross-section of each scatterering center (σ_b) and the distance of the point scatterer from the radar, as shown in

$$a_b^2 = \frac{P^{tx}G^{tx}G^{rx}\sigma_b\lambda_c^2}{(4\pi)^3 r_b^4}. \tag{7}$$

In (6) and (7), we have assumed stationary channel conditions and direct path target returns without any type of multipath. The RCS of a flat metallic triangular plate is a function of the radar aspect angle (θ_b), the plate area (A_b), and long dimension (d_b), [53], as shown in Fig. 4, and (8)

$$\sigma_b = \eta \frac{4\pi A_b^2 \cos^2\theta_b}{\lambda_c^2} \frac{\sin^4\left(\frac{2\pi}{\lambda_c}d_b\sin\theta_b\right)}{\left(\frac{2\pi}{\lambda_c}d_b\sin\theta_b\right)^4}. \tag{8}$$

The aspect angle is computed from the dot product of the incident vector
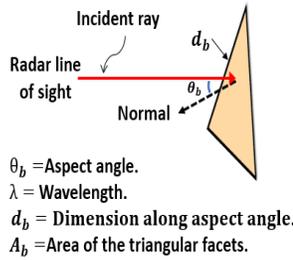
Fig. 4: Radar cross section of flat metallic triangular plate

from the radar to the plate and the normal vector of the plate. Since θ_b changes along the target trajectory, σ_b fluctuates. Not all scattering centers may be visible to the radar. One method for incorporating shadowing effects is to implement a shadowing algorithm on the radar data. However, since these algorithms utilize ray tracing methods, they are computationally expensive, especially for spatially large dynamic objects such as automobiles at millimeter-wave frequencies. For example, a ray-tracing algorithm for a 4mm wavelength would require the rays to be spaced one-tenth wavelength apart and emanate from an aperture that is at least the size of the largest automotive target that we consider. In our case, that would correspond to a truck of $8.5 \times 2.6m \times 5m$. Further, these computations would have to be repeated at the radar frame rate of several giga-Hertz as opposed to the video frame rates of computer animation algorithms typically used in vision. Instead, we opt for a low complexity algorithm by introducing a Bernoulli's random variable, η, to model a 20% visibility of each point scatterer [48]. The detailed comparison (accuracy and computational complexity) of the ray tracing and probalistic method is presented in the appendix.

The radar data is sampled at a frequency of $F_s = 1/\delta t$, and the fast time samples are numbered from $1 : N$. Similarly, if we consider a pth coherent processing interval (CPI) consisting of M PRIs, then the discrete representation of (6) is

$$S_b^{rx}[n,m] = a_b[m]\text{rect}\left[\frac{n-n_b}{N}\right]e^{-j\frac{4\pi f_c}{c}R_b}$$
$$e^{-j2\pi m f_{D_b}T_{PRI}}e^{j\pi K\delta t^2(n-n_b)^2} + \nu, \tag{9}$$

where n_b is the integer rounded from $\frac{\tau_b(t)}{\delta t}$.

We process the received signal using *stretch processing*, a variation of matched filtering, which is especially suited for low sampling frequencies [54]. The maximum unambiguous range of the radar, R_{max}, is equal to $\frac{cT_{PRI}}{2}$. For every CPI, we consider a radar range span of interest, $R_0 - \frac{R_{span}}{2} : R_0 + \frac{R_{span}}{2}$ within R_{max} where R_0 is called the central reference position (CRP). The time delay to the bth point scatterer can be expressed as $\tau_b = \tau_0 + \delta\tau_b(t)$ where $\tau_0 = \frac{2R_0}{c}$ corresponds to the time delay to the CRP. Since the target motion is known, the CRP is chosen to correspond to the mean range to $\vec{r}_C$ in every CPI. In stretch processing,

the received signal is multiplied with $e^{-j\pi K\delta t^2(n-n_0)^2}$, where n_0 is the integer rounded from $\frac{\tau_0}{\delta t}$ over every PRI. Thus, we obtain

$$S_b^{rx}[n,m] = a_b[m]e^{-j\frac{4\pi f_c}{c}R_b}e^{-\pi K\delta_t^2(n_0-n_b)^2}e^{-j2\pi m f_{D_b}T_{PRI}}$$
$$e^{-j4\pi K\delta_t^2 n(n_0-n_b)}. \tag{10}$$

We carefully compensate for the translational motion of the vehicle, and only consider the rotational motion of the point scatterer within a CPI. Then, the first two exponential terms in (10) are constant phase terms and are absorbed into the amplitude during further processing. The last two terms show the variation of the two-dimensional (2D) signal over slow and fast times, as shown in

$$S_b^{rx}[m,n] = a_b(\cdot)e^{-j2\pi f_{D_b}mT_{PRI}}e^{-j4\pi K\delta t^2 n(n_0-n_b)}. \tag{11}$$

The fast time sampling frequency ($F_s = 1/\delta t$) is obtained from twice the stretch bandwidth which is $\frac{2R_{span}K}{c}$ where R_{span} is much lower than R_{max}. Note that the stretch bandwidth is much lower than the radar bandwidth ($KT_{PRI} = \frac{2R_{max}K}{c}$). Hence, the sampling frequency requirements for stretch processing results in lower than that of the ordinary matched filtering.

When the target is an extended target with multiple point scatterers (B), then the received signal is obtained by the sum of the returns from each scatterer.

$$S^{rx}[m,n] = \sum_{b=1}^{B} S_b^{rx}[m,n]. \tag{12}$$

Here, we have ignored the multiple scattering between the different parts of the target. The output is processed using 2D Fourier transform to obtain range-Doppler ambiguity plots,

$$\chi[f_D, r] = \mathcal{DFT}_{2D}\{S^{rx}[m,n]\}, \tag{13}$$

where the range dimension r spans N steps from $R_0 - \frac{R_{span}}{2}$ to $R_0 + \frac{R_{span}}{2}$; and f_D spans M steps from $-\frac{1}{2T_{PRI}}$ to $\frac{1}{2T_{PRI}}$. The 2D plot can also be interpreted as a range-cross plot ($\chi[r, cr]$) provided an accurate estimate of the angular velocity (ω) of the target is available, since translational motion has been compensated. We estimate ω for every p^{th} CPI by the change in yaw (Θ) of the vehicle as shown in

$$\omega = \frac{\Theta[p] - \Theta[p-1]}{T_{CPI}}. \tag{14}$$

Then the Doppler axis is converted to the cross-range axis by

$$cr[m] = f_D[m] \times \frac{\lambda_c}{2\omega}, \text{for } m = 0 : M - 1. \tag{15}$$

Depending on ω, the cross-range spans across images may vary even when the pixel dimensions of the plot remain unchanged.

TABLE I: Automotive radar parameters for generating ISAR images

Parameters	Values
Carrier frequency ($2\pi f_c$)	77GHz
stretch Bandwidth ($\frac{1}{T_{SBW}}$)	8MHz
Sampling Frequency (F_s)	5MHz
Chirp rate (K)	60×10^{12} Hz^2
Chirp duration (T_{PRI})	$83.33\mu s$
Coherent processing interval (T_{CPI})	0.1s
Doppler resolution	10 Hz
Range resolution	0.075 m
Minimum cross-range resolution	0.19 m
Transmitted power (P_t)	25dBm

C. Noise and Range-Doppler Clutter Models

In this section, we discuss how we incorporated ground-based clutter along the range and wind-based clutter along the Doppler dimensions along with additive noise in the time-domain data.

Ground clutter: For a rough surface, the clutter cross-section is proportional to the surface clutter coefficient, σ^0, and the radar coverage area. A stable component - due to static road conditions such as road

material - and a fluctuating component due to wind contribute to σ^0 [41], [55]:

$$\sigma^0 = \sigma^0_{stable} + \sigma^0_{fluctuating}. \qquad (16)$$

We model σ^0 as an exponential random variable with a mean of -15dB which corresponds to asphalt at millimeter-wave frequencies [41], [56]. The radar coverage area is proportional to the antenna beamwidth (θ_{BW}), grazing angle (ψ), radar range resolution ($\delta r = \frac{c}{2KT_{PRI}}$) and range. Therefore, σ_c is

$$\sigma_c = \sigma^0 r \theta_{BW} \delta r \sec \psi. \qquad (17)$$

This results in a ground clutter that is a function of range as shown in

$$C_0[r] = \frac{P^{tx}G^{tx}G^{rx}\lambda_c^2 \sigma^0 \theta_{BW} \delta r \sec \psi}{(4\pi)^2 r^3}. \qquad (18)$$

We have maintained the radar at the height of 0.5m above the ground from which the grazing angle can be computed for every range r.

Doppler clutter: Based on [41], the power spectrum of the Doppler clutter can be modeled as a low pass filter response. When combined with the range related clutter, we obtain

$$C[f_D, r] = C_0[r] \left[1 + \left(\frac{f_D}{\Delta f_D} \right)^s \right]^{-1}. \qquad (19)$$

where s is a function of the average wind velocity (U) as shown in

$$s = \frac{2(U+2)}{(U+1)} \left(\frac{100}{2\pi f_c} \right)^{0.2}. \qquad (20)$$

In (19), Δf_D is the $-3dB$ width of the spectrum and is given by

$$\Delta f_D = 1.23 \left(\frac{3.2}{\lambda_c} \right) U^{1.3}. \qquad (21)$$

Based on local meteorological reports [57], U can vary from 0m/s to 10m/s. We consider four possible wind speeds (2.5, 5, 7.5, 10m/s) in our simulations. Finally, we convert the power values obtained from (19) to voltage values for each range-Doppler pixel ($c[f_D, r]$). We multiply the voltage with a phase modeled as a complex circularly symmetric random variable ($\phi[f_D, r]$). This complex clutter signal ($c[f_D, r]exp(j\phi[f_D, r])$) is then added to each pixel of the range-Doppler ISAR images $\chi[f_D, r]$ in (13). In this work, we use radar parameters that closely correspond to an actual automotive radar and list them in Table.I.

Noise: While clutter was modeled as a speckle noise in the radar images based on the above description, we modeled noise as an additive white Gaussian $\mathcal{N}(0, N_p)$ in the time-domain radar returns in (6). Based on our radar range equation in (7), the minimum received signal at the radar is -80 dBm. Based on the ratio between the minimum received signal at the radar and the mean noise floor, we considered four different signal to noise ratio (SNR) scenarios from $-5dB$ to $+10dB$ in our simulations.

In the next section, we experimentally demonstrate the validity of our simulation framework by comparing the results with those generated from measurement data.

D. Database of Simulated Data

We present examples of ISAR images of each of the targets below Figs 5, 6, 7, 8, and 9. For all five automotive targets (bicycle, auto-rickshaw, mid-size car, full-size car, and truck), we present two sets of results. The figures on the left show the ISAR images corrupted by additive noise in the receiver data due to receiver electronics. We present the set of images corresponding to an SNR of $+10dB$. The figures on the right show the ISAR images corrupted by range-Doppler clutter that give rise to speckle noise. The clutter strength, in this scenario, is a function of the surface reflection coefficient of the road and the wind speed. We present figures corresponding to a wind speed of 2.5m/s. In all of the figures, each row corresponds to images simulated for a distinct trajectory. The top row is obtained when the vehicle moves along an almost straight trajectory from N to S; the second row shows the trajectory of a target taking a left turn from E to S; the third row shows the trajectory of a target taking a right turn from S to E; while the fourth row shows the trajectory of a target

doing a U-turn from W to W. The range span (Y-axis) in all the figures is 20m and is centered along the CRP corresponding to the specific CPI. The cross-range span (X-axis) in all of the figures may vary from 10m to 20m and is centered at 0m. The cross-range axis correlates to the Doppler axis of $\chi[f_D, r]$. The noisy images on the left are of a dynamic range of $50dB$ from $-40dBm$ to $-90dBm$. While, the images on the right are of a dynamic range of 80dB from $-40dBm$ to $-120dBm$.

We, first, present the ISAR images of a *bicycle* in Fig.5. As mentioned earlier, the images in Fig.5a correspond to the ISAR images corrupted by noise, while the figures on the right correspond to the ISAR images corrupted by range-Doppler based clutter. The bicycle is a spatially narrow target and hence appears as a cluster of very closely spaced scattering centers, almost like a single-point scatterer. The dimensions of the target can be estimated from some of the images (for example, sub-figure vi). The noisy images on the left show that at some range positions, the target becomes difficult to discern due to low returns from the bicycle, especially when it is far from the radar. This is because of the low RCS of the bicycle. The cluttered images on the right show strong clutter at DC (corresponding to 0m along the cross-range). The width of the Doppler spectrum and the strength of the clutter returns change depending on the wind speed. The bicycle can still be discerned in some of the images along with the micro-Doppler tracks due to its wheels (sub-figures xi and xiv in (b)). Note that our model does not consider the motion of the bicycle handlebars or rider's body, and hence those effects cannot be observed in the images.

Next, we present the ISAR images of an *auto-rickshaw* in Fig.6. The images show that the auto-rickshaw is a spatially larger target than the bicycle. The shape of the vehicle is triangular in the top-view. In fact, in some of the top-view images (sub-figure vi), we can clearly see the dimensions of the vehicle. We also see considerable distortions along the Doppler (cross-range) dimension due to the micro-Doppler from the rotation of the wheels. Interestingly, in some images, we can see three distinct micro-Doppler tracks from the three wheels (sub-figure xi in (b)). On the top row, we observe that the longer dimension of the target's top view is oriented along the range dimension when the car moves from N to S. In the second row, the target is first oriented laterally and then turns length wise. This is because the target did a left turn from E to S. Similarly, in the third row, the target was first oriented along the long direction and then along the lateral direction as the target moved from S to E. Finally, in the last row, the target is always along the lateral direction since it does a U-turn from W to W. Therefore, the ISAR images offer some indication of the type of trajectory undertaken by a target.

Figure.7 presents the results of the mid-size car. Since this is a larger target than the auto-rickshaw, the returns are stronger. We are able to see the top view of the target with all four sides. Again, we observe some micro-Doppler based distortions along the cross-range due to the micro-Doppler returns from the four wheels. Four distinct micro-Doppler tracks are observed in the sub-figure x and xi in Fig.7b. Again, we observe the longer dimension of the car oriented along the range dimension when it is moving from either N to S or S to N. But the longer dimension of the car is oriented along the lateral dimension when the car is moving from E to W or vice versa.

The results in Fig.8, corresponding to the *full-size car*, look similar to those from the mid-size car in Fig.7, except for the larger dimensions of the car in the top-view. The dimensions of the full-size car are 5.7m $\times$ 2.4m whereas those of the mid-size car were 4.4m $\times$ 1.7m. The larger target also has stronger returns and is thus easily discerned in the images. Again, we are able to observe four distinct micro-Doppler tracks from the four wheels in some of the images (sub-figures ix-xii in Fig.8b). Also, we are able to see the changes in the orientation of the images as the car undergoes turns along its trajectory.

The largest automotive target that we have considered is the *four wheel truck*, for which the results are presented in Fig.9. Due to its large size, the top-view obtained from the ISAR images clearly presents the dimensions of the target, which are 8.5m $\times$ 2.6m. We are also able to observe the changes in the target orientation along the four distinct trajectories. The micro-Doppler distortions are considerably greater in this case due to the large wheel radii, and four micro-Doppler tracks in sub-figures xi in Fig.9 are well resolved in this case.
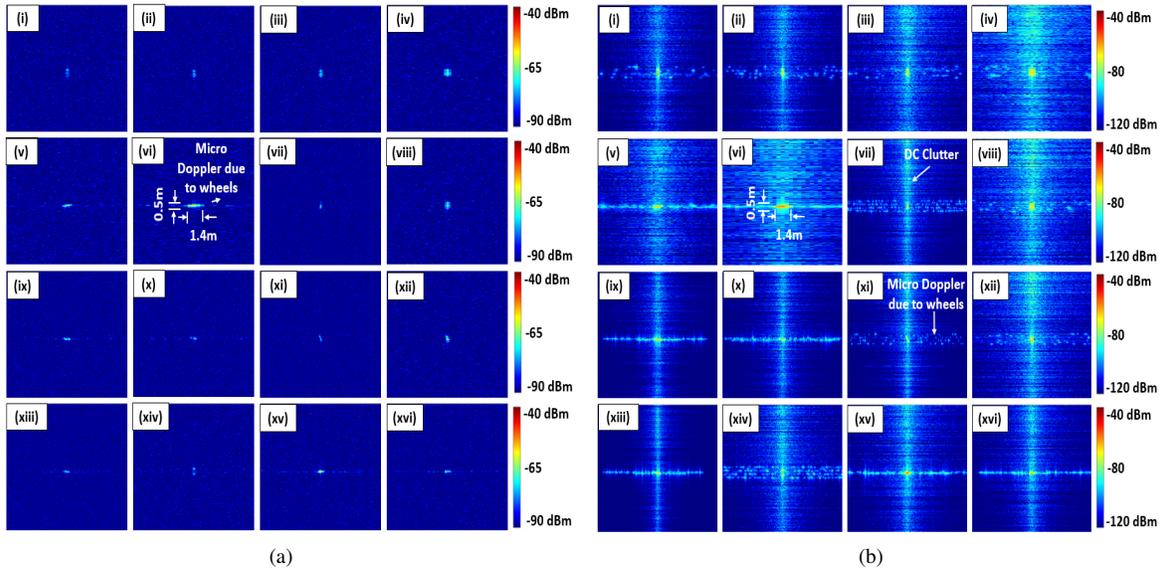
Fig. 5: ISAR images of a *bicycle* of (a) +10dB SNR, (b) with range-Doppler clutter with wind velocity 2.5 m/s at CPI frames corresponding to 1.5, 2.5, 3.5, 4.5s along following trajectories: (i-iv) straight path from north to south, (v-viii) left turn from east to south (ix-xii) right turn from west to south and (xiii-xvi) U-turn from west to west. The range span is 20m while the cross-range span varies from 10m to 20m. The dynamic range for SNR is 50dB (-40dBm to -90dBm) and for clutter 80dB (-40dBm to -120dBm).
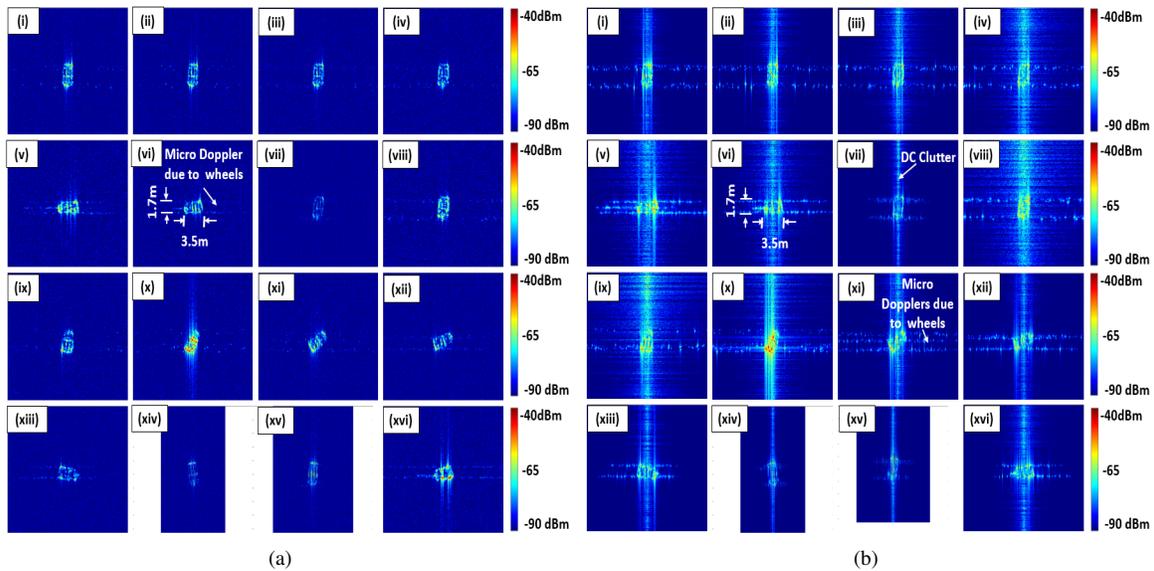
Fig. 6: ISAR images of an *auto-rickshaw* of (a) +10dB SNR (b) with range-Doppler clutter with wind velocity 2.5 m/s at CPI frames corresponding to 1, 2, 3, 4s along following trajectories: (i-iv) straight path from north to south, (v-viii) left turn from east to south (ix-xii) right turn from south to east and (xiii-xvi) U-turn from west to west. The range span is 20m while the cross-range span varies from 10m to 20m. The dynamic range for SNR is 50dB (-40dBm to -90dBm) and for clutter 80dB (-40dBm to -120dBm).

We list the complete set of simulated ISAR images in Table.II. To

TABLE II: Simulated ISAR image database

Type of images	Types of Targets (#)	Trajectories (#)	Images per trajectories	Total images (#)
Ideal Images	5	16	45-49	3750
Noisy images of SNR (+10,+5,0,-5 dB)	5	16	45-49	14976
Cluttered Images with wind velocities (2.5,5,7.5,10 m/s)	5	16	45-49	14976

summarize, we have considered five automotive targets - full-size car, mid-size car, truck, auto-rickshaw, and bicycle. Each target undergoes 16 trajectories, and each trajectory is of 5 seconds duration. Since each CPI is 0.1 seconds, we obtain between 45 and 49 images from each trajectory. The resulting range and minimum cross-range resolution in the ISAR images are 0.075×0.19m respectively. We obtain 3750 clean images that are free of noise and clutter. Then we corrupt these images with additive white Gaussian noise in the time-domain to obtain noisy images with SNR ranging from -5 to $+10$dB resulting in 14976 noisy images. Similarly, we introduce range-Doppler clutter with four different wind speeds (U) ranging from 2.5 m/s to 10 m/s to obtain 14976 cluttered images. With this paper, we publicly release this data set to the research community at https://tinyurl.com/msu6aj7yh.
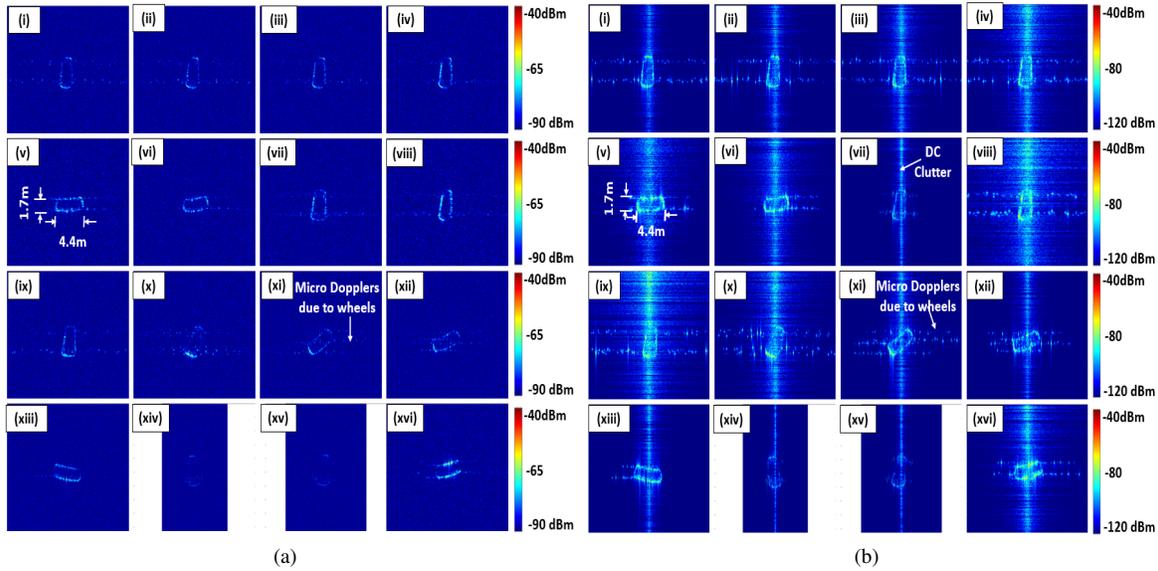
Fig. 7: ISAR images of *mid-size car* of (a) +10dB SNR, (b) with range-Doppler clutter with wind velocity 2.5 m/s at CPI frames corresponding to 1, 2, 3, 4s along following trajectories: (i-iv) straight path from north to south,(v-viii) left turn from east to south (ix-xii) right turn from south to east and (xiii-xvi) U-turn from west to west. The range span is 20m while the cross-range span varies from 10m to 20m. The dynamic range for SNR is 50dB (-40dBm to -90dBm) and for clutter 80dB (-40dBm to -120dBm).
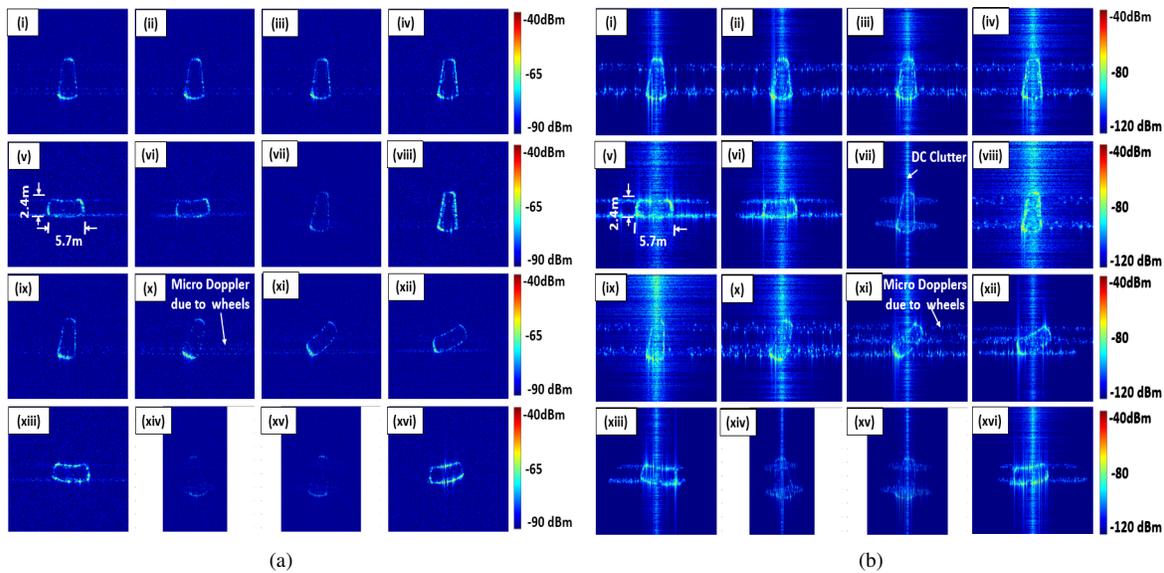
Fig. 8: ISAR images of *full-size car* of (a)+10dB SNR, (b) with range-Doppler clutter with wind velocity 2.5 m/s, at CPI frames corresponding to 1, 2, 3, 4s along following trajectories: (i-iv) straight path from north to south, (v-viii) left turn from east to south (ix-xii) right turn from south to east and (xiii-xvi) U-turn from west to west. The range span is 20m while the cross-range span varies from 10m to 20m.The dynamic range for SNR is 50dB (-40dBm to -90dBm) and for clutter 80dB (-40dBm to -120dBm).

III. MEASUREMENT DATA

The experimental set up for the measurements is shown in Fig.10. We configured the TI AWR-1843 radar to operate in an ultra-short range radar (USRR) mode and the corresponding parameters are listed in Table.I. The transmitted power from the radar is 14dBm, and the noise floor of the receiver is approximately -110dBm. We considered an auto-rickshaw (Fig.10a) of $2.6m \times 1.3m \times 1.7m$ dimensions, a bicycle (Fig.10b) of $1.6m \times 0.5m \times 1.5m$ dimensions, and a mid-sized car - Honda Brio (Fig.10c) of $3.6m \times 1.7m \times 1.5m$ dimensions for the measurement data collection. Note that these targets resemble the ones used in the simulations in terms of gross size. But the exact make and model are not the same in both the scenarios. We could not carry out experiments with the truck and large size car due to lack of availability of such targets in our premises. The simulation scenario considers 16 possible trajectories.

Not all of these trajectories could be replicated in our institute premises. Hence, we considered five trajectories as shown in Fig.11 - two straight trajectories, two right turn trajectories, and a left turn trajectory. These trajectories are similar to the trajectories chosen for the simulation data. In the two straight trajectories, the vehicles move along a straight path from left to right and right to left at an average speed of 15 kmph as shown in Fig.11b. In the first right turn trajectory, vehicles first move on the straight path and then execute a right turn in front of the radar, whereas in the second right turn trajectory, the vehicles first move on the straight path right to the radar and then execute a right turn as shown in Fig.11c at an average speed of 15 kmph. In the left turn trajectory, the vehicle first move on the straight path and then execute a left turn as shown in Fig.11d at an average speed of 15 kmph.

We performed matched filtering along the fast time and Doppler

(a)
(b)
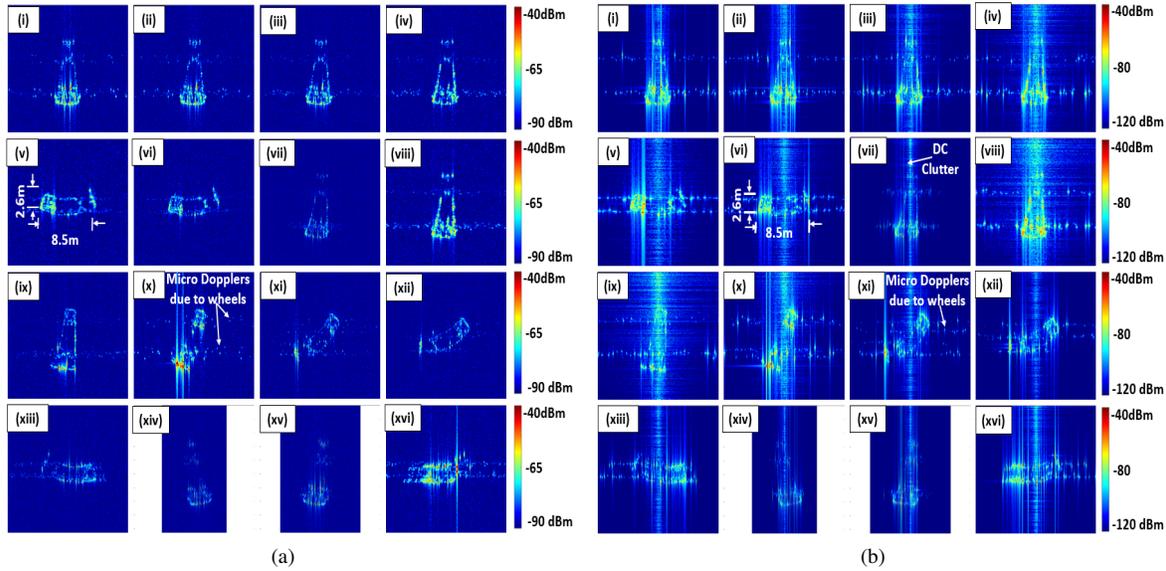
Fig. 9: ISAR images of *truck* of (a)+10dB SNR, (b) with range-Doppler clutter with wind velocity 2.5 m/s, at CPI frames corresponding to 1, 2, 3, 4s along following trajectories: (i-iv) straight path from north to south, (v-viii) left turn from east to south (ix-xii) right turn from south to east and (xiii-xvi) U-turn from west to west. The range span is 20m while the cross-range span varies from 10m to 20m. The dynamic range for SNR is 50dB (-40dBm to -90dBm) and for clutter 80dB (-40dBm to -120dBm).
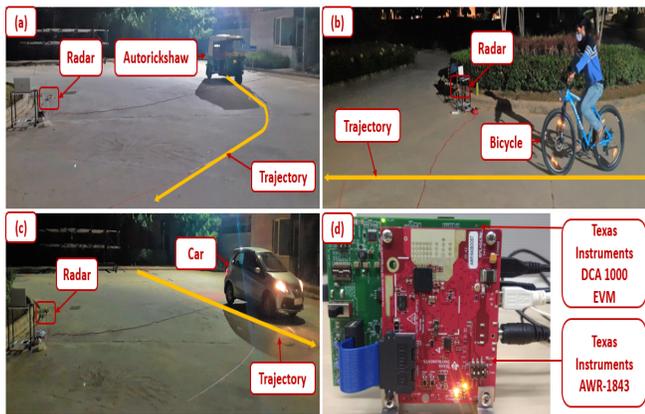
Fig. 10: Experimental setup for gathering ISAR images of (a) an auto-rickshaw undertaking a right turn trajectory, (b) a bicycle undertaking a straight trajectory, (c) a small sized car undertaking a left turn along their trajectory (d) Radar hardware.
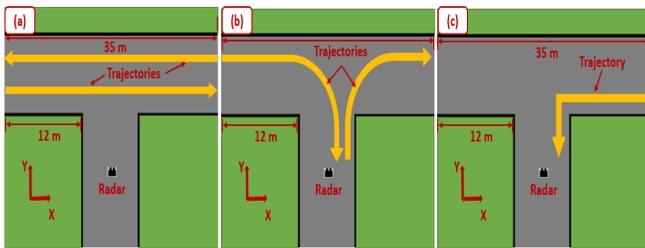
Fig. 11: Experimental setup and trajectories for gathering ISAR images of (a) two straight trajectories before the radar, (b) two right turn trajectories before the radar, (c) a left turn trajectory before the radar.

processing along the slow time to obtain ISAR images of the target for every CPI. We center the target in each measurement image by performing coarse and fine range compensation based on [58]. The radar data is carefully calibrated based on the radar range equation, and the dynamic range of the images are fixed from -40dBm to -120dBm. The ISAR images of a bicycle moving on left turn trajectory are shown in Fig.12 (i-iv) at CPI frame corresponding to 1,2,3,4 seconds. The Fig. 12(v-viii) are ISAR images of auto-rickshaw moving along the straight trajectory at CPI frame corresponding the 1,2,3,4 seconds. The Fig. 12(ix-xii) are the ISAR images of the car executing the straight trajectory at the CPI frame corresponding to 1,2,3,4 seconds. These images show the top-view dimensions of the vehicle along the range and cross-range dimensions from which we can infer the size of the vehicle. Interestingly, in the measurement results, we can observe the back of the car, which is not shadowed entirely by the front of the vehicles in some time intervals. We are able to see the obscured part of the vehicle due to the inherent advantages of ISAR processing, wherein reflections from different aspect angles get captured due to the relative rotational motion of the vehicle with respect to the radar. Another interesting observation is that the movie of the ISAR images provides insights on how the car is turning based on the orientation of the longer dimension of the car. In some frames, we observe a large spread along the cross-range dimension due to the micro-Dopplers arising due to the wheel motions. The strength of the radar returns are also fairly comparable between the two sets of results. We are also able to observe range based clutter in the measurement results increasing with for higher values of range as observed in the simulations. We also see Doppler-based clutter with the highest returns at cross-range values close to $0m$ corresponding to static and low-frequency clutter. Hence, there is, overall, a good qualitative agreement between the measured and simulated images. Quantitative comparison is not made since it was not possible to measure the exact clutter conditions in the real data or replicate the identical targets and trajectories in the two scenarios.

IV. CLASSIFICATION RESULTS

In this section, we use classical machine learning techniques - support vector machine (SVM) and random forests (RF) - and more recent deep learning algorithms based on transfer learning - Alexnet and Googlenet - for classifying the five automotive targets on the basis of their ISAR images. We will examine the impact of noise and clutter and the volume of test and training data on the classification performance. We will present
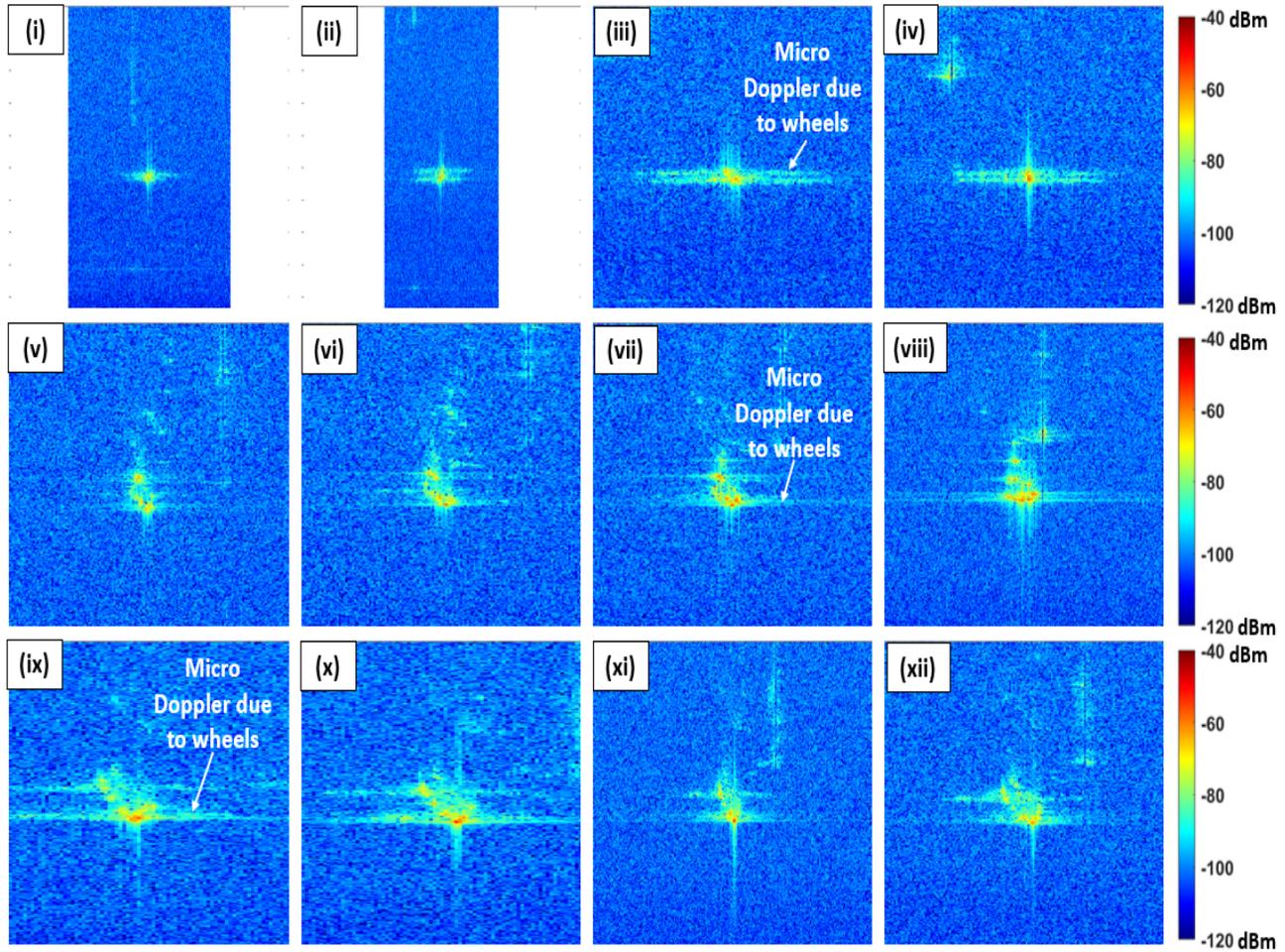
Fig. 12: ((i-iv) ISAR images of bicycle along the left turn trajectory at CPI frame corresponding to 1,2,3,4 s, (v-viii) ISAR images of auto-rickshaw along the straight trajectory at CPI frame corresponding to 1,2,3,4 s, and (ix-xii) ISAR images of car along the straight trajectory trajectory at CPI frame CPI frame corresponding to 1,2,3,4 s from measurement data. Range axis spans from 0 to 20m while the cross range axis spans -15 to +15m.

the classification result using measurement data ISAR images as test data for the classifier trained on the simulation data.

A. Effect of noise and clutter on classification performance

Based on Table.II, of the total volume of 14976 images for different SNR values, 70% are used for training and 30% for testing in the case of SVM and RF. In the case of Alexnet and Googlenet, we split the 30% data that are not used for training between validation and testing. The resulting classification accuracy for different SNR values is shown in Fig.13a. We first observe that the classification accuracy for all algorithms is above 75% even for low SNR of $-5dB$. The accuracy for SVM and RF are significantly poorer than those obtained from Alexnet and Googlenet at low SNR (-5dB). The classification accuracy improves for all cases as the SNR increases. The performances of Alexnet and Googlenet hold steady (above 80% for Googlenet and 90% for Alexnet) for all cases.

We perform a similar study where we examine the effect of clutter on the classification performance of the ISAR images. The clutter along the range is modeled using a mean surface clutter coefficient. As the range increases, the area of coverage increases resulting in greater clutter. Wind gives rise to Doppler based clutter along the cross-range dimension. Higher wind velocities (U) give rise to greater clutter. Again, we have assumed a 70-30 split between training and test data for SVM and RF and a 70-15-15 split between training, validation, and test for Alexnet and Googlenet. We show the variation of the classification accuracy with respect to mean wind velocity in Fig.13b. We observe that the classification performance for all the algorithms is fairly high (above 85%) even for high values of clutter arising from high wind speeds
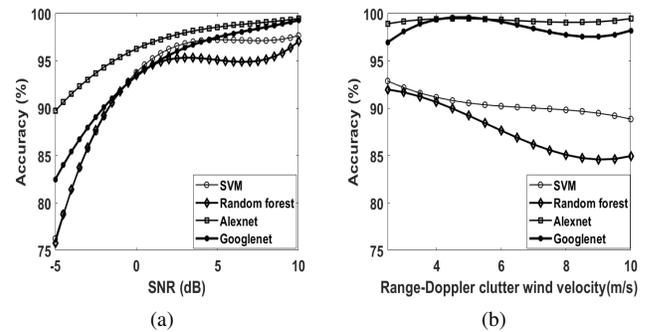
Fig. 13: Classification accuracy of automotive target ISAR images using SVM, random Forest and transfer learning algorithms based on Alexnet and Googlenet for (a) differing SNR values and (b) different range-Doppler clutter values. 70% data used for training and remaining data used for validation and testing.

(10m/s). The performance of the two transfer learning-based algorithms (Alexnet and Googlenet) remains consistent even for the high values of clutter. On the other hand, we observe a slight deterioration in the performance of the two classical machine learning techniques with higher clutter values. The results indicate that the ISAR images offer highly discriminatory features for classification, even in the presence of high noise and clutter.

B. Effect of test and training percentages

In Table.III, we report the classification results for the four algorithms - SVM, RF, Alexnet, and Googlenet - for different percentages of training, testing, and validation data. For each case, we have performed a 5-fold cross-validation. We first consider the data that are just corrupted by noise (SNR varying from $+10dB$ to $-5dB$). Then we repeat the tests on data that are just corrupted by only clutter (wind velocities varying from 2.5 to 10m/s). Finally, we repeat the tests on data that are corrupted by both noise and clutter. We observe that the performances for all the cases are above 87%. Since, the training data is very large; the algorithms perform well even when the training and testing are split evenly. We do not see a significant improvement in the performance with an increase in training data. The transfer learning algorithms like Alexnet and Googlenet perform very well, even for low SNR and high clutter.

In the following section, we present the confusion matrices obtained from the classification of data combining both noise and clutter. These results are obtained using 70% training data. For each case, the rows show the true labels of the data, while columns show the labels of the predictions. We present three metrics with each confusion matrix. They are *precision*, *recall* and the F_1 score. Precision is defined as the ratio of the true positives (the highlighted number along the diagonal) over the sum of the true positives and false positives (column sum); while recall is defined as the ratio of the true positives over the sum of the true positives and false negatives (row sum). For each case, we also provide the F_1 score, which is defined as the harmonic mean of average precision and average recall,

$$F_1 = 2\frac{avg.precision \times avg.recall}{avg.precision + avg.recall}. \tag{22}$$

The first confusion matrix is presented for the SVM in Table.IV. We observe the least confusion for the truck. Due to its large dimensions and strong returns, the truck is rarely mistaken for any of the other targets or vice versa. Similarly, the bicycle is very small and hence not easily confused with the other targets. However, due to its weak returns and small size, sometimes, the bicycle is not easily discernible in noisy images. The two cars are often confused by each other due to the similarity in their dimensions, the number of wheels, and their strength of returns. The mid-size car, especially, has the lowest precision and recall because it is most similar to both the full-size car and the auto-rickshaw. The F_1 score for SVM is 88.6%.

We observe a similar result for the RF classifier in the second matrix in Table.IV. Again, the best precision and recall are observed for the large truck and the small bicycle. The results of the bicycle are slightly worse than the truck because of its weak returns, which get affected when the noise floor is high. Again, the two cars are confused by each other. However, this time, the results for the full-size car have significantly improved while those of the mid-size car have only slightly improved. There is a noticeable improvement in the performance of the Alexnet classifier, compared to the traditional machine learning algorithms for all cases, as reported in the third matrix in Table.IV. Here, both the precision and recall for all the cases are above 95%. Thus the two cars are no longer significantly confused by each other. The confusion between the auto-rickshaw and mid-size car has also substantially decreased. The same improvement is also observed for the Googlenet classifier as seen in the fourth matrix in Table.IV. Again, the accuracy is above 95% for all five cases, both precision and recall. The F_1 scores for RF, Alexnet, and Googlenet are 92%, 98.1%, and 97.6%, respectively. Note that in all of the cases discussed above, we use only a single target instance per class. The limitation is due to the limited availability of open data of computer animated models of vehicles provided by their manufacturers. The diversity is incorporated by varying the trajectories, speeds, noise and clutter conditions in the radar data. More robust classification tests with multiple instances from the same class will be presented in future studies.

C. Testing measured data with classifiers trained with simulation data

We present the classification results of two classifiers - SVM based on classical machine learning and Alexnet transfer learning-based deep learning technique. Both classifiers are trained on simulation data. We use 99 images for each of the three automotive targets - bicycle, auto-rickshaw and medium-sized car - as test images for the classifier performance evaluation. In Table.V we present confusion matrices for both the classifiers. For the SVM classifier, we observe that the least confusion is for the bicycle because it is smaller (in terms of spatial extent) than the other targets. Confusion arises between the car and auto-rickshaw because of their comparable sizes. The performance of the Alexnet classifier is better than the SVM classifier and the confusion between the auto-rickshaw and the mid-size car is reduced. The F_1 scores for SVM and Alexnet are 75.3 %, and 94.2% respectively. The results indicate that the simulated ISAR images are of high fidelity and can be used to train classifiers that can subsequently be used on real test data.

V. CONCLUSION

We have demonstrated an automotive radar simulation framework that incorporates radar scattering phenomenology of commonly found road vehicles as well as range-based surface clutter and Doppler-based wind clutter and additive receiver noise. Using this simulation framework, we have demonstrated that high-resolution ISAR radar images, characterized by the fine range and cross-range resolution of dynamic automotive targets, can be generated with millimeter-wave automotive radars. A large database of over 30000 images has been publicly released to the radar community. The simulation framework has been verified through experimental data gathered with a real automotive millimeter-wave radar from Texas Instruments. These images provide meaningful information about the dimensions of the vehicle along the top-view as well as the number of wheels and the trajectory undertaken by the vehicle in the case of larger vehicles such as auto-rickshaws, cars, and trucks. Smaller targets such as bicycles, on the other hand, more closely resemble single-point scatterers. These images indicate the robustness of ISAR images as automotive target classification features for both traditional machine learning techniques as well as the more recent deep neural networks.Due to restrictions in the data collection facilities within the institute premises, we have limited our data collection to single target instances of each target class. This may be satisfactory in the cases of auto-rickshaw where the models and makes are fairly standard. However, in other types of vehicles, there may be greater diversity in each target class. The testing of the classifiers on multiple instances of the same target class will form the focus of future works.

Also, note that the specific problem addressed in the paper falls in the category of supervised learning - where the classifiers are trained with data from specific targets and tested with data from the same set of targets. In real world scenarios - especially in automotive radar settings - we may encounter new targets that may have previously not been seen during training. Such problems would fall under the category of unsupervised learning. The exploration of new and better algorithms for handling the unknown radar data will form the basis of our future work.

VI. APPENDIX

A. Study of effects of shadowing on radar signatures

In our simulation model, we resolve the target into triangular facets and identify a scattering center at the centroid of each of the facets. Then we sum the returns from all the facets to compute the time-domain radar returns. Naturally, some of these facets would not be visible to the radar due to shadowing. The visibility of a facet can be determined through ray tracing. However, the process is computationally complex and scales both with the carrier frequency and sampling frequency of the radar data. To compare the computational complexity, the ray tracing algorithm (based on the surface normals of the facets) was implemented on a Intel(R) Xeon(R) CPU E5-2620 V3 processor running at 2.40 GHz using 16 GB of RAM. The duration of data collection for a single frame (shown in Fig.14(vi)) was 179.33 minutes. The computation time of ray tracing algorithm using advanced GPU reduces significantly, but the cost of the system increases which makes the technique challenging for generating large volumes of data [45]. Instead, we incorporate a Bernoulli's random variable (η) to account for the visibility. The computational time for generating the ISAR image using the probabilistic

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIV.2022.3146639, IEEE Transactions on Intelligent Vehicles

11

Classifier	Training (%)	Testing (%)	Validation (%)	SNR	Range-Doppler Clutter	Combined
SVM	70	30	-	92.4	93.8	88.6
	60	40	-	92.2	93.7	88.1
	50	50	-	92.3	93.3	87.7
Random Forest	70	30	-	90.9	93.3	91.9
	60	40	-	90.6	93.2	91.9
	50	50	-	90.9	92	90.8
Alexnet	70	15	15	96.7	99.9	98.1
	60	20	20	96.4	99.4	98.1
	50	25	25	95.5	97.7	97.3
Googlenet	70	15	15	95.9	99.2	97.6
	60	20	20	94.7	99.2	97.3
	50	25	25	94.4	98.6	97.2

TABLE III: Classification of ISAR images using classical machine learning algorithms - SVM and random forest, and deep learning based algorithms - Alexnet and Googlenet

SVM		Predicted Labels					
	Vehicle	Auto-rickshaw	Bicycle	Full size car	Mid size car	Truck	Recall
True Labels	Auto-rickshaw	1655	3	41	96	2	92.1
	Bicycle	24	1687	17	72	3	93.6
	Full size car	48	39	1478	193	34	82.5
	Mid size car	111	112	146	1463	2	79.8
	Truck	14	2	56	15	1673	95.1
	Precision	89.4	91.5	85.0	79.6	97.6	
RF		Predicted Labels					
	Vehicle	Auto-rickshaw	Bicycle	Full size car	Mid size car	Truck	Recall
True Labels	Auto-rickshaw	1674	14	13	96	0	93.2
	Bicycle	4	1723	11	63	2	95.6
	Full size car	29	34	1577	145	7	88.0
	Mid size car	65	130	67	1570	2	85.6
	Truck	3	4	24	18	1711	97.2
	Precision	94.3	90.4	93.2	83.0	99.4	
Alexnet		Predicted Labels					
	Vehicle	Auto-rickshaw	Bicycle	Full size car	Mid size car	Truck	Recall
True Labels	Auto-rickshaw	895	1	0	2	0	99.7
	Bicycle	0	897	1	1	0	99.8
	Full size car	3	1	876	14	5	97.4
	Mid size car	5	40	5	849	0	94.4
	Truck	2	0	4	0	893	99.3
	Precision	98.9	95.5	98.9	98.0	99.4	
Googlenet		Predicted Labels					
	Vehicle	Auto-rickshaw	Bicycle	Full size car	Mid size car	Truck	Recall
True Labels	Auto-rickshaw	881	0	8	7	2	98.1
	Bicycle	0	890	0	9	0	99.0
	Full size car	0	3	863	32	1	96.0
	Mid size car	0	37	2	860	0	95.7
	Truck	0	0	8	1	891	99.0
	Precision	100	95.7	98.0	94.6	99.7	

TABLE IV: Confusion matrices from SVM, RF, Alexnet and GoogleNet classifiers based on 70% training, 15% validation and 15% test data.

SVM		Predicted Labels					
	Vehicle	Auto-rickshaw	Bicycle	Full size car	Mid size car	Truck	Recall
True Labels	Auto-rickshaw	76	22	0	2	0	76.0
	Bicycle	4	92	0	4	0	93.9
	Full size car	0	0	0	0	0	-
	Mid size car	22	25	1	50	2	51.0
	Truck	0	0	0	0	0	-
	Precision	75.5	66.2	-	89.3	-	
Alexnet		Predicted Labels					
	Vehicle	Auto-rickshaw	Bicycle	Full size car	Mid size car	Truck	Recall
True Labels	Auto-rickshaw	80	0	0	18	1	80.8
	Bicycle	0	99	0	0	0	100
	Full size car	0	0	0	0	0	-
	Mid size car	0	0	0	99	0	100
	Truck	0	0	0	0	0	-
	Precision	100	100	-	84.6	-	

TABLE V: Confusion matrices from SVM and Alexnet and GoogleNet classifiers, trained with simulation data and test with measured data.

method was a few seconds. To study the effectiveness/accuracy of this probabilistic approach, we compare the ISAR signature results with that obtained from ray tracing in Fig.14. The first five subfigures show the
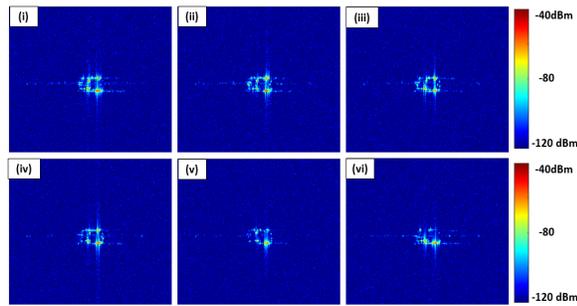
Fig. 14: (i-v) ISAR images of auto-rickshaw for the same CPI frame using Bernoulli's random variable with visibility of 60 %, 50%, 40 %, 30%, 20 % respectively. (vi) is the ray tracing algorithm ISAR images of auto rickshaw for the same CPI frame. The range axis spans from 0 to 20m while the cross-range axis spans -15 to +15m.

ISAR images of the auto-rickshaw with visibility of η fixed at 60%, 50%, 40%, 30%, and 20% respectively, and (vi) is the ISAR image of the same frame using ray tracing. We observe in the result from ray tracing, that the nearer portions of the target are more distinctive than the distant portions though they are somewhat visible. Visually, the image with 20% visibility resembles the ray tracing result most closely. Quantitatively, we compare the results using two metrics - the structural similarity index measure (SSIM) and the normalized mean square error (NMSE) - and present it in Table.VI. The SSIM compares gross features like luminosity and contrast between two images while NMSE does pixel-wise comparison. The table

TABLE VI: Structural similarity index measure (SSIM) in ISAR images generated using ray tracing and Bernoulli random variable

% of facets visible	SSIM	NSME
60	1	0.8636
50	1	0.4470
40	1	0.4337
30	1	0.3465
20	1	0.3411

shows that the SSIM is high for all five cases but the NMSE is lowest for a visibility of 20%. Further, in the ISAR images generated using the measurement data in Fig.12, we observe that some of the obscured parts of the vehicle are not fully shadowed. Due to these reasons, we use the approximate and less computationally expensive probabilistic method, with a visibility coefficient of 20%, to account for the shadowing effect in place of the ray tracing algorithm.

REFERENCES

[1] R. H. Rasshofer and K. Gresser, "Automotive radar and lidar systems for next generation driver assistance functions." *Advances in Radio Science*, vol. 3, 2005.

[2] M. Schneider, "Automotive radar-status and trends," in *German microwave conference*, 2005, pp. 144–147.

[3] J. Hasch, E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-wave technology for automotive radar sensors in the 77 ghz frequency band," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 3, pp. 845–860, 2012.

[4] B. Fleming, "Recent advancement in automotive radar systems [automotive electronics]," *IEEE Vehicular Technology Magazine*, vol. 7, no. 1, pp. 4–9, 2012.

[5] Y. Cui, R. Chen, W. Chu, L. Chen, D. Tian, Y. Li, and D. Cao, "Deep learning for image and point cloud fusion in autonomous driving: A review," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–18, 2021.

[6] I. Jouny, F. Garber, and S. Ahalt, "Classification of radar targets using synthetic neural networks," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 2, pp. 336–344, 1993.

[7] H.-C. Chiang, R. L. Moses, and L. C. Potter, "Model-based classification of radar images," *IEEE Transactions on Information Theory*, vol. 46, no. 5, pp. 1842–1854, 2000.

[8] I. Bilik, J. Tabrikian, and A. Cohen, "Gmm-based target classification for ground surveillance doppler radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 267–278, 2006.

[9] I. Urazghildiiev, R. Ragnarsson, P. Ridderstrom, A. Rydberg, E. Ojefors, K. Wallin, P. Enochsson, M. Ericson, and G. Lofqvist, "Vehicle classification based on the radar measurement of height profiles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 245–253, 2007.

[10] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Tracking of extended objects with high-resolution doppler radar," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 5, pp. 1341–1353, 2016.

[11] S. Kuutti, R. Bowden, Y. Jin, P. Barber, and S. Fallah, "A survey of deep learning applications to autonomous vehicle control," *IEEE Transactions on Intelligent Transportation Systems*, 2020.

[12] X. Cai, M. Giallorenzo, and K. Sarabandi, "Machine learning-based target classification for mmw radar in autonomous driving," *IEEE Transactions on Intelligent Vehicles*, pp. 1–1, 2021.

[13] Y. Kim and H. Ling, "Human activity classification based on micro-doppler signatures using a support vector machine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1328–1337, 2009.

[14] Y. Kim and T. Moon, "Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 8–12, 2015.

[15] G. Diraco, A. Leone, and P. Siciliano, "A radar-based smart sensor for unobtrusive elderly monitoring in ambient assisted living applications," *Biosensors*, vol. 7, no. 4, p. 55, 2017.

[16] S. Vishwakarma and S. S. Ram, "Dictionary learning with low computational complexity for classification of human micro-dopplers across multiple carrier frequencies," *IEEE Access*, vol. 6, pp. 29 793–29 805, 2018.

[17] M. S. Seyfioğlu, A. M. Özbayoğlu, and S. Z. Gürbüz, "Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1709–1723, 2018.

[18] J. Le Kernec, F. Fioranelli, C. Ding, H. Zhao, L. Sun, H. Hong, J. Lorandel, and O. Romain, "Radar signal processing for sensing in assisted living: The challenges associated with real-time implementation of emerging algorithms," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 29–41, 2019.

[19] F. Fioranelli, M. Ritchie, and H. Griffiths, "Classification of unarmed/armed personnel using the netrad multistatic radar for micro-doppler and singular value decomposition features," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 9, pp. 1933–1937, 2015.

[20] R. Prophet, M. Hoffmann, A. Ossowska, W. Malik, C. Sturm, and M. Vossiek, "Pedestrian classification for 79 ghz automotive radar systems," in *2018 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2018, pp. 1265–1270.

[21] P. Khomchuk, I. Stainvas, and I. Bilik, "Pedestrian motion direction estimation using simulated automotive mimo radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 3, pp. 1132–1145, 2016.

[22] S. Lee, Y.-J. Yoon, J.-E. Lee, and S.-C. Kim, "Human-vehicle classification using feature-based svm in 77-ghz automotive fmcw radar," *IET Radar, Sonar & Navigation*, vol. 11, no. 10, pp. 1589–1596, 2017.

[23] M. Ritchie, F. Fioranelli, H. Borrion, and H. Griffiths, "Multistatic micro-doppler radar feature extraction for classification of unloaded/loaded micro-drones," *IET Radar, Sonar & Navigation*, vol. 11, no. 1, pp. 116–124, 2016.

[24] P. Molchanov, K. Egiazarian, J. Astola, R. Harmanny, and J. De Wit, "Classification of small uavs and birds by micro-doppler signatures," in *2013 European Radar Conference*. IEEE, 2013, pp. 172–175.

[25] I. Bilik and J. Tabrikian, "Radar target classification using doppler signatures of human locomotion models," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 4, pp. 1510–1522, 2007.

[26] S. Chen, H. Wang, F. Xu, and Y.-Q. Jin, "Target classification using the deep convolutional networks for sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4806–4817, 2016.

[27] S. A. Wagner, "Sar atr by a combination of convolutional neural network and support vector machines," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 6, pp. 2861–2872, 2016.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIV.2022.3146639, IEEE Transactions on Intelligent Vehicles

13

[28] Z. Lin, K. Ji, M. Kang, X. Leng, and H. Zou, "Deep convolutional highway unit network for sar target classification with limited labeled training data," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 7, pp. 1091–1095, 2017.

[29] S. S. Ram and A. Majumdar, "High-resolution radar imaging of moving humans using doppler processing and compressed sensing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 2, pp. 1279–1287, 2015.

[30] V. C. Chen, *Inverse Synthetic Aperture Radar Imaging; Principles*. Institution of Engineering and Technology, 2014.

[31] M. Martorella, E. Giusti, A. Capria, F. Berizzi, and B. Bates, "Automatic target recognition by means of polarimetric isar images and neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3786–3794, 2009.

[32] M. Vespe, C. Baker, and H. Griffiths, "Automatic target recognition using multi-diversity radar," *IET Radar, Sonar & Navigation*, vol. 1, no. 6, pp. 470–478, 2007.

[33] S.-H. Park, M.-G. Joo, and K.-T. Kim, "Construction of isar training database for automatic target recognition," *Journal of Electromagnetic Waves and Applications*, vol. 25, no. 11-12, pp. 1493–1503, 2011.

[34] R. Paladini, M. Martorella, and F. Berizzi, "Classification of man-made targets via invariant coherency-matrix eigenvector decomposition of polarimetric sar/isar images," *IEEE transactions on geoscience and remote sensing*, vol. 49, no. 8, pp. 3022–3034, 2011.

[35] A. A. Danylov, T. M. Goyette, J. Waldman, M. J. Coulombe, A. J. Gatesman, R. H. Giles, X. Qian, N. Chandrayan, S. Vangala, K. Termkoa *et al.*, "Terahertz inverse synthetic aperture radar (isar) imaging with a quantum cascade laser transmitter," *Optics express*, vol. 18, no. 15, pp. 16 264–16 272, 2010.

[36] J. S. Kulpa, M. Malanowski, D. Gromek, P. Samczyński, K. Kulpa, and A. Gromek, "Experimental results of high-resolution isar imaging of ground-moving vehicles with a stationary fmcw radar," *International Journal of Electronics and Telecommunications*, vol. 59, no. 3, pp. 293–299, 2013.

[37] C. Li and H. Ling, "Wide-angle, ultra-wideband isar imaging of vehicles and drones," *Sensors*, vol. 18, no. 10, p. 3311, 2018.

[38] H. Essen, M. Hagelen, W. Johannes, R. Sommer, A. Wahlen, M. Schlechtweg, and A. Tessmann, "High resolution millimetre wave measurement radars for ground based sar and isar imaging," in *2008 IEEE Radar Conference*. IEEE, 2008, pp. 1–5.

[39] N. Pandey, G. Duggal, and S. S. Ram, "Database of simulated inverse synthetic aperture radar images for short range automotive radar," in *2020 IEEE International Radar Conference (RADAR)*. IEEE, 2020, pp. 238–243.

[40] M. I. Skolnik *et al.*, *Introduction to radar systems*. McGraw-hill New York, 1980, vol. 3.

[41] G. P. Kulemin, *Millimeter-wave radar targets and clutter*. Artech House, 2003.

[42] J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters*, vol. 9, no. 3, pp. 293–300, 1999.

[43] A. Liaw, M. Wiener *et al.*, "Classification and regression by randomforest," *R news*, vol. 2, no. 3, pp. 18–22, 2002.

[44] L. Torrey and J. Shavlik, "Transfer learning," in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*. IGI global, 2010, pp. 242–264.

[45] Y. Deep, P. Held, S. S. Ram, D. Steinhauser, A. Gupta, F. Gruson, A. Koch, and A. Roy, "Radar cross-sections of pedestrians at automotive radar frequencies using ray tracing and point scatterer modelling," *IET Radar, Sonar & Navigation*, vol. 14, no. 6, pp. 833–844, 2020.

[46] S. S. Ram, C. Christianson, Y. Kim, and H. Ling, "Simulation and analysis of human micro-dopplers in through-wall environments," *IEEE Transactions on Geoscience and remote sensing*, vol. 48, no. 4, pp. 2015–2023, 2010.

[47] G. Duggal, K. V. Mishra, and S. S. Ram, "Micro-doppler and micro-range detection via doppler-resilient 802.11ad-based vehicle-to-pedestrian radar," in *IEEE Radar Conference 2019*, 2019, pp. 1–6.

[48] G. Duggal, S. Vishwakarma, K. V. Mishra, and S. S. Ram, "Doppler-resilient 802.11ad-based ultra-short range automotive joint radar-communications system," *IEEE Transactions on Aerospace and Electronic Systems*, 2020.

[49] J. Lekner, *Theory of reflection of electromagnetic and particle waves*. Springer Science & Business Media, 1987, vol. 3.

[50] J. R. Wait, *Electromagnetic waves in stratified media: Revised edition including supplemented material*. Elsevier, 2013, vol. 3.

[51] C. Li, Z. Peng, T.-Y. Huang, T. Fan, F.-K. Wang, T.-S. Horng, J.-M. Munoz-Ferreras, R. Gomez-Garcia, L. Ran, and J. Lin, "A review on recent progress of portable short-range noncontact microwave radar systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 5, pp. 1692–1706, 2017.

[52] S. S. Ram, "Representation of radar micro-dopplers using customized dictionaries," *Topics in Radar Signal Processing*, p. 197, 2018.

[53] G. T. Ruck, D. E. Barrick, W. D. Stuart, and C. K. Krichbaum, *Radar cross section handbook*. Plenum press New York, 1970, vol. 1.

[54] M. A. Richards, *Fundamentals of radar signal processing*. Tata McGraw-Hill Education, 2005.

[55] F. Ulaby, M. C. Dobson, and J. L. Álvarez-Pérez, *Handbook of radar scattering statistics for terrain*. Artech House, 2019.

[56] E. S. Li and K. Sarabandi, "Low grazing incidence millimeter-wave scattering models and measurements for various road surfaces," *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 5, pp. 851–861, 1999.

[57] P. Pandey, D. Kumar, A. Prakash, J. Masih, M. Singh, S. Kumar, V. K. Jain, and K. Kumar, "A study of urban heat island and its association with particulate matter during winter months over delhi," *Science of the Total Environment*, vol. 414, pp. 494–507, 2012.

[58] B. Haywood and R. Evans, "Motion compensation for isar imaging," in *Proceedings of Australian Symposium on Signal processing and applications*, 1989, pp. 112–117.

Shobha Sundar Ram completed her Masters and Ph.D. from the University of Texas at Austin, USA in Electrical and Computer Engineering in 2006 and 2009 respectively. She is currently an Associate Professor at the Indraprastha Institute of Information Technology (IIIT) Delhi. She serves as a Associate Editor for the IEEE Transactions in Aerospace and Electronics Systems. Her research interests are in electromagnetics and radar signal processing with a focus on through-wall radar and automotive radar.

Neeraj Pandey is pursuing his Ph.D. degree in electrical engineering from Indraprastha Institute of Information Technology Delhi. He obtained his Master's in Microwave Electronics from the University of Delhi, South Campus, India, in 2018. His research interests are in automotive radar, radar signal processing, and machine learning.