Azimuth Super-Resolution for FMCW Radar in Autonomous Driving

Yu-Jhe Li[1] Shawn Hunt[2] Jinhyung Park[1] Matthew O'Toole[1] Kris Kitani[1]
[1]Carnegie Mellon University [2]DENSO International America, Inc.
{yujheli, jinhyun1, motoole2, kmkitani}@andrew.cmu.edu, {shawn.hunt}@na.denso.com

Abstract

We tackle the task of Azimuth (angular dimension) super-resolution for Frequency Modulated Continuous Wave (FMCW) multiple-input multiple-output (MIMO) radar. FMCW MIMO radar is widely used in autonomous driving alongside Lidar and RGB cameras. However, compared to Lidar, MIMO radar is usually of low resolution due to hardware size restrictions. For example, achieving 1° azimuth resolution requires at least 100 receivers, but a single MIMO device usually supports at most 12 receivers. Having limitations on the number of receivers is problematic since a high-resolution measurement of azimuth angle is essential for estimating the location and velocity of objects. To improve the azimuth resolution of MIMO radar, we propose a light, yet efficient, Analog-to-Digital super-resolution model (ADC-SR) that predicts or hallucinates additional radar signals using signals from only a few receivers. Compared with the baseline models that are applied to processed radar Range-Azimuth-Doppler (RAD) maps, we show that our ADC-SR method that processes raw ADC signals achieves comparable performance with 98% (50 times) fewer parameters. We also propose a hybrid super-resolution model (Hybrid-SR) combining our ADC-SR with a standard RAD super-resolution model, and show that performance can be improved by a large margin. Experiments on our Pitt-Radar dataset and the RADIal dataset validate the importance of leveraging raw radar ADC signals. To assess the value of our super-resolution model for autonomous driving, we also perform object detection on the results of our super-resolution model and find that our super-resolution model improves detection performance by around 4% in mAP. The Pitt-Radar and the code will be released at the link.

1. Introduction

We address the task of azimuth angle super-resolution for Frequency Modulated Continuous Wave (FMCW) [25] Multiple Input Multiple Output (MIMO) [26] radar in autonomous driving. In addition to Lidar and RGB cameras, Radar has been commonly used for autonomous vehicles [4, 5] due to its robustness in adverse weather conditions (*e.g.,* fog, snow, rain) with longer wavelength. FMCW MIMO radar uses a line array of receiver antennas to capture reflected signals of multiple chirps sent out by a line array of transmitters. We can characterize the location (orientation and distance) and velocity of nearby objects from a radar's Range-Azimuth-Doppler (RAD) map. The RAD map is computed using the Discrete Time Fast Fourier Transform (FFT) [11, 26] on the discretized receiver signals after Analog-to-Digital conversion (ADC). In the domain of autonomous driving, MIMO [26] radar devices typically have low resolution due to the physical constraint on the size of the sensors. For example, a device with eight antenna receivers has at most an angle resolution of about 15° [11]. Thus, it is important to develop new technologies which can increase the azimuth resolution of radar sensing,

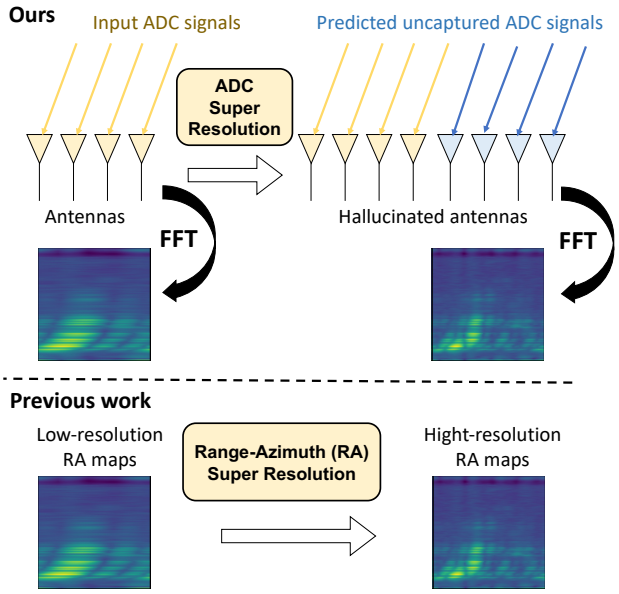

Figure 1. Previous work perform super-resolution from low-resolution and high-resolution Range-Azimuth maps after applying Discrete Time Fourier Transform (FFT). Instead, we aim to predict more ADC uncaptured signals before they are transformed into Range-Azimuth (RA) maps.

without requiring a large sensor size.

For MIMO radar [26], both range resolution and velocity (Doppler) resolution can be improved by setting different bandwidths (the range of frequencies used for signal pulses) and frame times (duration of a single signal pulse). However, azimuth resolution is strictly dependent on the radar sensor's hardware specifications, such as antenna size, and cannot be improved by changing the parameters of the radar sensor [15, 34], which restricts radar imaging performance. To improve the azimuth resolution of radar Range-Azimuth (RA) maps, some work proposes to use regularization approaches from signal processing [31, 32, 35] and deep models [2, 10] for azimuth super-resolution. However, such prior work is limited to processing the RA maps (not even using Doppler information) as shown in Figure 1. Since RAD and RA maps are the transformed data after keeping the magnitude only after FFTs, the information loss regarding the relationship between each receiver may lead to limited performance for azimuth super-resolution. In addition, how to enhance the azimuth resolution from processing ADC signals also remains challenging.

In order to tackle the challenge of azimuth angle super-resolution on the ADC signals, we propose a light, yet efficient, ADC super-resolution (ADC-SR) model for azimuth resolution. Since the azimuth resolution is related to the number of receivers each capturing independent signals, we aim to predict uncaptured signals from hallucinated receivers. For example, the model takes ADC signals recorded with 4 receivers and outputs the predicted ADC signals with 8 receivers, as shown in Figure 1. To the best of our knowledge, our ADC-SR is the first to apply deep models to ADC signals for azimuth super-resolution. In addition, the hallucinated ADC signals can then be processed with FFTs [26] to obtain high-resolution RAD maps for further use in autonomous driving such as object detection. We note that we can further refine the RAD maps with the RAD super-resolution (RAD-SR) model. To evaluate and compare our approach with other baseline models, we have collected a dataset named Pitt-Radar which contains ADC signals. Compared with RAD-SR which relies on RAD data, our ADC-SR achieves comparable performance with fewer network parameters. We also show that our hybrid pipeline named Hybrid-SR combining ADC-SR and RAD-SR improves the baseline model by a large margin. Moreover, since naive bilinear downsampling of the RAD map does not truly reflect the outputs of lower-resolution radar sensors, we propose a more theoretically grounded way of downsampling for training later.

To assess the value of our super-resolution model for autonomous driving, we also evaluate the performance of existing object detectors trained along with our super-resolution models. The improvements in the experiments demonstrate our approach is applicable to downstream tasks in autonomous driving. The contributions of this paper can be summarized as follows:

- We propose an azimuth super-resolution model named ADC-SR which takes into complex ADC radar signals and predicts the signals from unseen receivers, which is able to produce higher-resolution RAD maps.
- We propose a hybrid model named Hybrid-SR for improved performance on RAD-SR when combined with our ADC-SR. To make comparisons with all of the baseline models, a downsampling method is also proposed for evaluation.
- We propose a MIMO radar dataset named Pitt-Radar which contains ADC signals for benchmarking.
- Our developed model not only achieves satisfactory azimuth super-resolution on our collected Pitt-Radar and one benchmark dataset but also improves the existing object detector by a large margin.

2. Related Works

Radar azimuth super-resolution. To improve the azimuth resolution in Range-Azimuth maps, many works leveraging signal processing techniques have been proposed. Since the azimuth signal can be modeled as a convolution of a target distribution and an antenna pattern [9, 33], the azimuth resolution can be improved by deconvolutions. Several deconvolution methods have been proposed to such as Wiener filtering (WF) [7], truncated singular value decomposition (TSVD) [8], Tikhonov regularization (REGU) [8], Richardson–Lucy (RL) [6], and iterative adaptive approach (IAA) [20]. Sparse regularization has been proposed as another line of methods by solving a regularization problem using split Bregman algorithm (SBA) [31, 32, 35]. While these non-deep models only achieve limited performance, recently some deep methods are also proposed to address the azimuth super-resolution in Range-Azimuth maps. Armanious *et. al.* [2] propose an adversarial network for super-resolution on micro-Doppler imagery. Geiss *et. al.* [10] apply Unet on the super-resolution of weather radar maps. However, Range-Azimuth maps are transformed with FFTs ADC signals and few works have been explored for the super-resolution directly on the ADC signals.

Radar for autonomous driving. Radar is resilient to fog, rain, and snow, but one drawback to radar is that it produces low-resolution images, which makes it very challenging for the tasks of object recognition or semantic segmentation in autonomous driving. Current automotive datasets, such as NuScenes [5] and RADDet [30], contain radar data using technology that relies on MIMO radar. Some works for object detection on these datasets also are proposed in [21, 27, 28]. While MIMO configurations are able to derive informative velocity features using multiple sent chirp

Table 1. Publicly available Radar datasets for object detection in autonomous driving. We only list commonly used datasets.

Radar Device	Dataset	FOV	Range	Azimuth	Doppler	Point Cloud	ADC
Scaning radar	Oxford Radar RoboCar [3]	360	✓	✓	✗	✗	✗
	RADIATE [24]	360	✓	✓	✗	✗	✗
	MulRan [12]	360	✓	✓	✗	✗	✗
MIMO radar	CRUW [29]	<180	✓	✓	✓	✗	✗
	Nuscenes [5]	<180	✗	✗	✓	✓	✗
	Zendar [17]	<180	✓	✓	✓	✗	✗
	CARRADA [18]	<180	✓	✓	✓	✗	✗
	RADDet [30]	<180	✓	✓	✓	✗	✗
	RadarScenes [23]	<180	✗	✗	✓	✓	✗
	RADIal [21]	<180	✓	✓	✓	✓	✓
	Pitt-Radar (Ours)	<180	✓	✓	✓	✓	✓

signals, it has low azimuth resolution. On the other hand, some datasets such as Oxford Radar RoboCar [3] and RADIATE [24] are collected with 360° scanning radar using a Navtech [1] device. This scanning device measures each azimuth using a moving antenna and provides as high of an azimuth resolution as Lidar. Some works for object detection using scanning radar are proposed in [14, 19]. Yet, scanning radar currently does not provide Doppler features as MIMO radar. In this work, we focus on applying azimuth super-resolution specifically for MIMO radar which has all features of the range, azimuth, and Doppler.

3. Radar Datasets

3.1. Current datasets

Currently, in autonomous driving, there are two types of Frequency Modulated Continuous Wave (FMCW) [25] radar devices used: 1) scanning radar and 2) MIMO radar. Scanning radar usually employs the Navtech [1] radar, which provides 360° high-resolution range-azimuth maps. However, scanning radar does not provide the Doppler feature as MIMO radar which uses several transmitters and receivers to measure the range, azimuth, and Doppler information through FFTs. Compared with scanning radar, MIMO has low azimuth resolution in RAD maps. The overview of the current datasets is summarized in Table 1.

Scanning Radar. There are currently few available datasets providing radar data from scanning radar such as Oxford Radar RoboCar [3], RADIATE [24], and MulRan [12]. Though these datasets are of high resolution and have 360° field of view (FOV), they do not have the Doppler feature or even raw ADC signals. In addition, they only provide data in bird-eye-view Range-Azimuth 2D format.

MIMO Radar. There are many available autonomous driving datasets that provide MIMO radar, as shown in Table 1. As we can see from the table, all of these datasets using MIMO radar provide a Doppler feature and most of them provide range-azimuth-Doppler maps. However, only RADlal [21] and our collected dataset provide additional ADC signals for use.

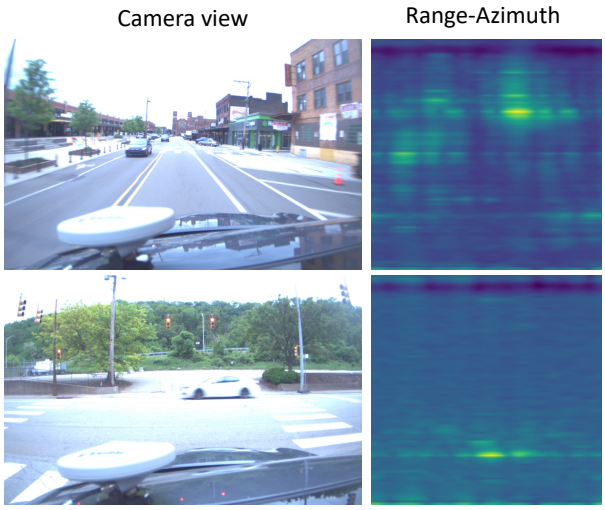

Figure 2. Samples of Pitt-Radar.

3.2. Pitt-Radar

In order to train and evaluate the super-resolution models trained on ADC radar signals, we collect a dataset named Pitt-Radar in Pittsburgh, PA, USA. The data were collected on a Lexus SUV that was equipped with stereo cameras, a LIDAR, a high-precision GPS device, and a Texas Instruments AWR1843BOOST radar sensor. The stereo cameras were logged at 15 Hz, the LIDAR was logged at 10 Hz, the GPS was logged at 25 Hz, and the radar was logged at 30 Hz. The log contains high-precision timestamps and was synchronized offline. The Texas Instrument's radar exposes many parameters in their mmWave Studio software package. The parameters include operating frequency, idle time, frequency slope, number of ADC samples, and sampling rate. We present the examples of the dataset in Figure 2 and more details for the experimental settings of the dataset can be obtained later in the experiment. We note that the difference between our dataset and RADlal [21] lies in we have multiple device setups (*e.g.,* 128, 256, and 512 range samples) while RADlal [21] only provides data with one setup.

4. The Proposed Approach

4.1. Problem Formulation and Overview

Given ADC radar signals from few receivers, we aim to predict unseen ADC radar signals which are able to produce a higher resolution of RAD maps. Specifically, we are given ADC radar signals $I_{ADC} \in \mathbb{C}^{N_{rec} \times N_{chirp} \times N_{adc}}$ from N_{rec} receivers, where each receiver receives signals of N_{chirp} chirps of N_{adc} ADC discrete complex samples. Our goal is to produce super-resolved ADC radar signals I^{SR}_{adc}, which is close to I^{HR}_{adc} with more N^{HR}_{rec} receivers, from I^{LR}_{adc} with few N^{LR}_{rec} receivers. For example, the high-resolution ADC signals I^{HR}_{adc} are recorded from 4 receivers and 3 transmitters while low-resolution data I^{LR}_{adc} are recorded from 1 or 2 transmitters. In practice, the field of view (FOV) of all the ADC data is identical. In addition, all raw ADC data I_{adc} can be transformed into Range-Azimuth-Doppler (RAD) maps $I_{rad} \in \mathbb{C}^{N_{Range} \times N_{Azimuth} \times N_{Doppler}}$ through Fast Fourier Transforms (FFTs) for downstream tasks such as object detection.

The overview of our proposed pipeline: Hybrid-SR is presented in Figure 3, which is composed of two separate yet sequential stages: 1) ADC super-resolution (ADC-SR) and 2) RAD super-resolution (RAD-SR). First, in the ADC-SR stage, we introduce an ADC-ConvNet with S upsample factor which takes I^{LR}_{adc} as input and produces the intermediate feature $F^{SR}_{adc} \in \mathbb{R}^{2*S*N_{rec} \times N_{chirp} \times N_{adc}}$. Then the feature F^{SR}_{adc} is transformed to I^{SR}_{adc} with the feature pixel shuffling module. The hallucinated ADC signals I^{SR}_{adc} can be processed into informative RAD I^{SR}_{rad} maps with FFTs. Second, in order to refine the super-resolved RAD maps I^{SR}_{rad}, we introduce a RAD-Unet which takes into I^{SR}_{rad} and produce the refined I^{SSR}_{rad} in the RAD-SR stage. During inference, we can either apply Hybrid-SR, ADC-SR, or RAD-SR independently for either the input formats of ADC signals or RAD data.

4.2. Downsampling by removing transmitters

In order to train the super-resolution models, we need to prepare the high-resolution and low-resolution pairs. Specifically, we sample the ADC pair I^{HR}_{adc}, I^{LR}_{adc} or RAD pair I^{HR}_{rad}, I^{LR}_{rad} for training. Theoretically, the field of view (FOV) of RAD pair I^{HR}_{rad}, I^{LR}_{rad} needs be matched so the range-azimuth maps can be aligned spatially. Thus, we maintain the same FOV when downsampling the high-resolution data by removing the antennas correctly from I^{HR}_{adc} to produce I^{LR}_{adc}.

Since the FOV depends on the distance between each receiver (density of the receivers) [11], the density needs to be the same after removing the receivers. We now present a preliminary introduction of virtual antenna receivers [11] as shown in Figure 4 and how we maintain FOV. If we are given 4 receivers uniformly spaced with distance d and 3 transmitters in Figure 4a the signals we observe from the 4 receivers can be rearranged to signals from 12 virtual receivers with distance d using one transmitter, shown in Figure 4b. Since the time of sending signals for the 3 transmitters are different (see Figure 4c), the 4 physical receivers are able to distinguish which kinds of signals they receive. Therefore, the direct way to downsample the I^{HR}_{adc} correctly to produce I^{LR}_{adc} and maintain density is to remove the side transmitters, *e.g.,* Tx 1 or Tx3. Then, the RA maps we obtain from I^{HR}_{rad} and I^{LR}_{rad} will also have the same FOV and are spatially aligned (see the input and ground truths in Figure 6).

4.3. Azimuth super-resolution from ADC signals

In order to produce I^{SR}_{adc} close to I^{HR}_{adc}, we propose the ADC super-resolution (ADC-SR) for learning the mapping between I^{LR}_{adc} and I^{HR}_{adc}. Unlike traditional super-resolution tasks in image super-resolution [13], super-resolution of radar data is more complicated since it has 1) complex values and 2) predicting missing signals from unseen receivers is different from classical pixel upsampling.

To address the first issue with complex values, we separate the real and imaginary parts into two channels. Specifically, $I_{adc} \in \mathbb{C}^{N_{rec} \times N_{chirp} \times N_{adc}}$ will be transformed to $I^{'}_{adc} \in \mathbb{R}^{2 \times N_{rec} \times N_{chirp} \times N_{adc}}$ with two channels. We then introduce ADC-SRNet denoted as E_{ADC} to produce super-resolved feature maps: $F^{SR}_{adc} = E_{ADC}(I^{LR}_{adc}) \in \mathbb{R}^{2*S*N_{rec} \times N_{chirp} \times N_{adc}}$, where S indicates the upsampling factor. To produce the correct format of I^{SR}_{adc}, we introduce feature pixel shuffling function g to reshape the intermediate features into I^{SR}_{adc}:

$$I^{SR}_{adc} = g(E_{ADC}(I^{LR}_{adc})) \in \mathbb{C}^{S*N_{rec} \times N_{chirp} \times N_{adc}}, \quad (1)$$

which has S times more receivers than I^{LR}_{ADC}. In order to train the ADC-SRNet (E_{ADC}), we employ the ADC loss using L2 distance between low-resolution and high-resolution ADC signal pairs:

$$\mathcal{L}_{ADC} = \mathcal{L}_2(g(E_{ADC}(I^{LR}_{adc})), I^{HR}_{adc}), \quad (2)$$

4.4. Azimuth super-resolution from RAD maps

In order to additionally refine the radar data with RAD format, we propose a baseline model for RAD super-resolution (RAD-SR) which learns the mapping between I^{LR}_{rad} and I^{HR}_{rad}. Both of the RAD data can be obtained by applying Fast Fourier Transforms (FFTs) [11] on I^{LR}_{adc} and I^{HR}_{adc}. While RAD data may contain complex values, we usually only keep the magnitude of the RAD data for radar object detection [30]. Thus, we can derive the RAD magnitude radar maps as $I_{rad} = \|FFTs(I_{adc})\| \in \mathbb{R}^{N_{Range} \times N_{Azimuth} \times N_{Doppler}}$. Since both I^{LR}_{rad} and I^{HR}_{rad} have the same 3D shape, the super-resolution task in RAD can be formulated as 3D image-to-image translation task.

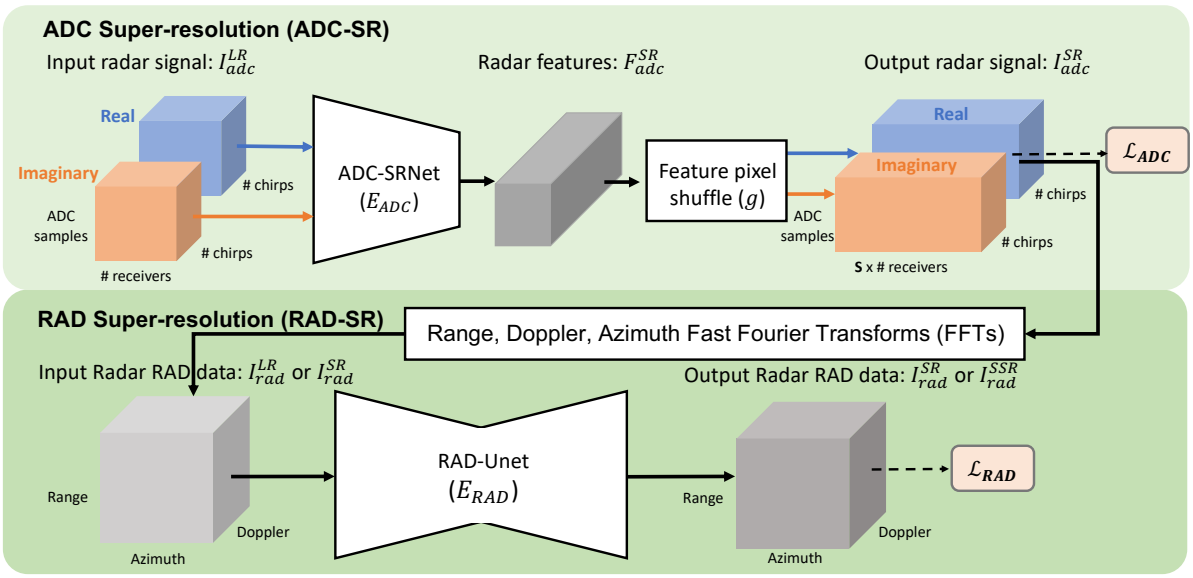

Figure 3. **Overview of our proposed hybrid super-resolution (Hybrid-SR) model.** Our Hybrid-SR has two streams: ADC-SR and RAD-SR. In the ADC-SR stage, we introduce an ADC-ConvNet with S upsample factor which takes I_{adc}^{LR} as input and produces intermediate features F_{adc}^{SR}, which are then super-resolved I_{adc}^{SR} using the 3D pixel shuffling module. The ADC data is then processed into informative RAD I_{rad}^{SR} maps with FFTs. Then, our RAD-Unet takes I_{rad}^{SR} and produces the finalized I_{rad}^{SSR} in the RAD-SR stage.

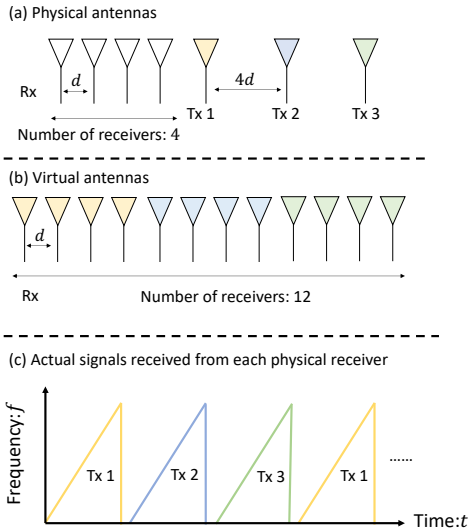

Figure 4. Illustration of virtual antennas with receivers (Rx) and transmitters (Tx). The signals produced by 4 receivers and 3 transmitters can be transformed virtually into 12 receivers.

Therefore, we employ RAD-Unet (E_{RAD}) employing Unet with 3D convolutions to produce I_{rad}^{SR} by taking into I_{rad}^{LR} as input:

$$I_{rad}^{SR} = E_{RAD}(I_{rad}^{LR}) \in \mathbb{R}^{N_{Range} \times N_{Azimuth} \times N_{Doppler}}. \quad (3)$$

To train the RAD-Unet, we utilize the RAD loss using L2 distance between low-resolution and high-resolution RAD data:

$$\mathcal{L}_{RAD} = \mathcal{L}_2(E_{RAD}(I_{rad}^{LR}), I_{rad}^{HR}), \quad (4)$$

We note that traditional non-deep upsampling methods (*i.e.,* interpolation or extrapolation) or deep methods [13] may not work since the shapes of I_{RAD}^{LR} and I_{RAD}^{HR} are identical in terms of the pixel resolution. As we see the input and ground truth in Figure 6, the main difference lies in the size and shape of the local responses inside the Range-Azimuth maps.

While we have presented two super-resolution approaches ADC-SR and RAD-SR, both ADC-SR and RAD-SR can be combined sequentially as our final model: Hybrid-SR. Specifically, the input of RAD-Unet (E_{RAD}) would be the RAD maps generated ADC-SR and we can derive the output RAD maps as:

$$I_{rad}^{SSR} = E_{RAD}(\|FFTs(I_{adc}^{SR})\|), \quad (5)$$

where

$$I_{rad}^{SR} = \|FFTs(I_{adc}^{SR})\| \in \mathbb{R}^{N_{Range} \times N_{Azimuth} \times N_{Doppler}}. \quad (6)$$

We can train the entire pipeline using the total hybrid loss:

$$\mathcal{L}_{Hybrid} = \mathcal{L}_{ADC} + \mathcal{L}_{RAD}. \quad (7)$$

Though both ADC-SR and RAD-SR are two independent streams to perform super-resolution, ADC-SR is assumed to be more effective since it leverages the ADC signals before FFTs are applied, which is also supported later in the experiments. Because of FFTs, RAD radar I_{rad} will always have the same shape $N_{Range} \times N_{Azimuth} \times N_{Doppler}$ regardless of how many receivers the RAD maps are produced from. This will also make super-resolution

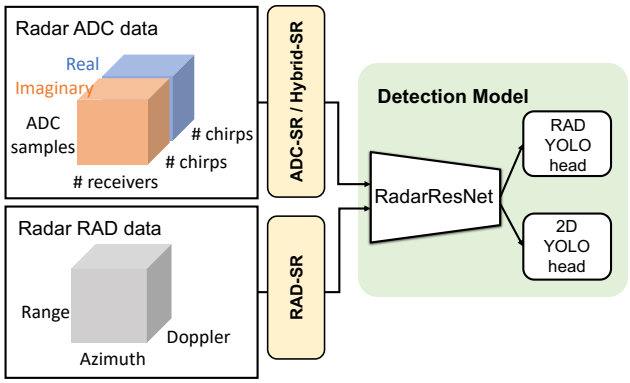

Figure 5. The proposed pipeline for adapting our azimuth super-resolution model to the existing detector, *i.e.*, RADDet [30].

harder for RAD if the model is deployed to unknown receivers in the real world. Thus, our new proposed ADC-SR is more applicable.

4.5. Adapting our model to object detectors

In order to leverage our proposed models for autonomous driving, we present the pipeline for adapting azimuth super-resolution to an object detector using RADDet [30] in Figure 5. For either ADC signals or RAD data, we apply the corresponding azimuth super-resolution model. Specifically, we apply ADC-SR or Hybrid-SR on the ADC signals and RAD-SR on the post-processed RAD data. We would like to note that, our azimuth super-resolution models can be treated as a plug-and-play tool for fine-tuning the existing models. Later in the experiments, we will also show that our models can even directly bring the performance gain without fine-tuning the detectors.

5. Experiments

5.1. Datasets

We use Pitt-Radar and RADIal [21] for training and evaluating our super-resolution models. On the other hand, we use RADDet [30] for evaluating the performance of object detection when combining radar super-resolution models with one existing detector proposed in [30].

Pitt-Radar. Pitt-Radar contains around 10,000 ADC radar frames in total with 4 receivers and 3 transmitters. We use the first half of the 5,000 frames for training and the second half of the 5,000 frames for testing.

RADIal. RADIal [21] contains 91 sequences for a total of 2 hours using 4 radar devices each has 3 transmitters and 4 receivers. Each sequence contains raw ADC signals recorded with their native frame rate. There are approximately 25,000 radar frames. We only use the first 5,000 frames for training the super-resolution models and the second 5,000 frames for testing.

RADDet. RADDet [30] provides a total number of 10,158 radar frames with 8 virtual receivers and only has data with RAD format. This dataset provides the 3D RAD annotations and 2D bird-eye-view annotations. It also contains the pre-selected train/test split where 80% of the radar frames are for training.

5.2. Settings and evaluation protocols

Azimuth super-resolution. To evaluate our super-resolution models, we use the low-resolution data with the middle 4 receivers and high-resolution radar data with all 12 receivers as ground truths. We evaluate the performance of the 3D RAD map reconstruction and 2D RA map reconstruction. Specifically, given different formats of input data for each model, *e.g.,* ADC signals for ADC-SR and RAD data for RAD-SR, we can always produce the same output format: RAD maps with FFTs. We can also further derive the RA maps by taking summation over the Doppler dimension. Thus, we can compute the performance of reconstruction using the metrics of mean square error (MSE) and peak signal-to-noise ratio (PSNR).

Object detection. To test the performance of object detection, we apply our super-resolution model on the RadarResnet detector proposed by RADDet [30]. Specifically, we also utilize their both RAD YOLO head and 2D YOLO head for predicting different types of bounding box locations. Following RADDet [30], we evaluate the model using mean average precision (mAP) in the COCO evaluation [16] with different IOU: 0.1, 0.3, 0.5, and 0.7 for a fair comparison.

5.3. Implementation details

We implemented our model using PyTorch. Following the device setup in RADIal [21] and RADDet [30], the number of chirps N_{chirp} is 256. The number of ADC samples N_{adc} in signals is 256 for Pitt-Radar and RADDet, and 512 for RADIal due to device settings. While we have recorded our Pitt-Radar using different settings of ADC samples, *e.g.,* 512, 256, or 128, which can lead to different range resolutions, we use the ones identical to RADDet. The number of receivers N_{rec} is set as 4 for low-resolution and set as 12 for high-resolution during training. For RAD radar FFT process, the range sample size N_{range} is the same as the number of ADC samples N_{adc} via range-FFT. The number of Doppler samples $N_{Doppler}$ is the same as N_{chirp} via Doppler-FFT. The azimuth size $N_{Azimuth}$ is set as 256 with zero paddings before applying azimuth-FFT. More details for how FFTs work to derive range, azimuth, and Doppler information can be obtained in [26]. For ADC-SRNet, we employ 6 residual blocks using 3D convolutions and skip-connections similar to SRGAN [13]. For RAD-Unet, we employ 3D convolutions following 2D Unet [22]. The difference lies in the upsampling blocks inside the RAD-Unet

Table 2. Comparisons with baseline models on the radar azimuth super-resolution by evaluating reconstruction of RA maps.

Method	Input	Pitt-Radar		RADlal	
		MSE ↓	PSNR ↑	MSE ↓	PSNR ↑
2D SRGAN	RA	21.4	23.2	25.7	25.2
2D UNet	RA	12.5	30.1	11.3	29.6
3D SRGAN	RAD	17.6	28.2	18.5	27.3
RAD-SR	RAD	5.8	33.9	4.4	35.1
ADC-SR	ADC	6.2	33.1	5.1	34.6
Hybrid-SR	ADC	**0.9**	**39.6**	**1.0**	**38.9**

Table 3. Comparisons with baseline models on the radar azimuth super-resolution by evaluating reconstruction of RAD cubes.

Method	Input	Pitt-Radar		RADlal	
		MSE ↓	PSNR ↑	MSE ↓	PSNR ↑
3D SRGAN	RAD	10.2	17.1	9.5	18.4
RAD-SR	RAD	2.2	24.2	2.5	23.8
ADC-SR	ADC	3.9	21.1	4.2	20.2
Hybrid-SR	ADC	**0.8**	**31.5**	**0.9**	**30.9**

utilizing the trilinear upsampling method. We use the Adam optimizer with a learning rate of 0.001 throughout all of the experiments. The batch size is set as 4 on a single NVIDIA 2080 Ti GPU.

5.4. Results and comparisons in super-resolution

RA maps. To evaluate the azimuth super-resolution and evaluate the effectiveness of different input types of radar data, we compare our proposed azimuth super-resolution models: ADC-SR, RAD-SR, and Hybrid-SR with some 2D baseline models: SRGAN [13], 2D UNet [22], and 3D baseline model: 3D SRGAN (with 3D convolutions). 2D baseline models will take into the low-resolution 2D Range-Azimuth (RA) as inputs. We present the reconstruction results of Range-Azimuth (RA) 2D maps in Table 2. First, we observe that existing strong 2D image super-resolution models such as SRGAN or Unet fail to conduct super-resolution for Range-Azimuth maps. This indicates that the 2D RA maps processed from 3D RAD maps may lead to information loss which is essential for super-resolution. Second, although 3D SRGAN takes as input RAD maps, it still fails to reconstruct 2D RA maps compared with RAD-SR. We credit the reason to that because the task of producing high-resolution RA maps from low-resolution RAD maps can not be formulated as the task of pixel-upsampling, while this task is rather like image-to-image translation. Thus, RAD-SR employing 3D Unet works better for reconstructing high-resolution RA maps from RAD maps. Third, ADC-SR achieves comparable performance as RAD-SR, and our Hybrid-SR combining ADC-SR exhibits superior reconstruction performance by taking into the ADC signals. This infers that performing super-resolution using ADC signals can improve the reconstruction result.

Table 4. Comparison of network parameters

Method	3D SRGAN	RAD-SR	ADC-SR	Hybrid-SR
Params.	90k	2512k	66k	2578k

RAD maps. Since RA maps are the post-processed data by removing the Doppler feature, we further evaluate the performance of reconstruction of entire 3D RAD maps. To evaluate the RAD reconstruction, we also compare our models (ADC-SR, RAD-SR, and Hybrid-SR) with 3D SRGAN while existing 2D models can not produce 3D data. We summarize the results in Table 3. First, we can still observe the same situation for 3D SRGAN which fails to reconstruct the RAD maps. Second, our proposed ADC-SR in Hybrid-SR taking into ADC signals outperforms RAD-SR which only takes RAD maps as inputs for azimuth super-resolution by a large margin. This also infers that performing super-resolution using raw ADC signals can have better reconstruction.

Network parameters. To evaluate the efficiency of all of our proposed azimuth super-resolution models, we present the comparisons of the number of network parameters in Table 4. We can observe that our ADC-SR is the lightest model among the four compared models while it has comparable performance in azimuth super-resolution as RAD-SR. This demonstrates that ADC-SR is able to perform efficient azimuth super-resolution when taking into the raw ADC signals compared with RAD maps.

5.5. Results and comparisons in object detection

In order to assess the effectiveness of our proposed azimuth super-resolution models for autonomous driving tasks, we evaluate the performance gains brought by our models on a previous radar object detector, *i.e.,* Radar-ResNet in RADDet [30]. There are two types of YOLO heads used in RADDet: RAD YOLO head for producing 3D bounding boxes in RAD maps and 2D YOLO head for 2D boxes in RA maps.

RAD YOLO head. We evaluate the RADDet detector using different kinds of super-resolution models: RAD-SR, ADC-SR, and Hybrid-SR when producing the RAD bounding boxes. We would like to note that, we can still apply inverse Fast Fourier Transform (IFFT) on RAD data for deriving the original ADC signals if the number of receivers is known even if RADDet does not originally provide ADC signals. We found that discrete IFFT will be correct numerically by filtering out the tiny values (thresholding with value $1e-6$). We present the result of detection using RAD YOLO head in Table 5. We report the performances both on fixing and fine-tuning the detector plugging in our trained super-resolution models. First, we observe that directly

Table 5. Results of radar object detection using RAD YOLO head following RADDet [30]

Method	$AP_{0.1}$	$AP_{0.3}$	$AP_{0.5}$	$AP_{0.7}$
RADDet	0.764	0.563	0.251	0.059
RADDet (fix.) + RAD-SR	0.761	0.596	0.273	0.061
RADDet (fix.) + ADC-SR	0.765	0.582	0.262	0.063
RADDet (fix.) + Hybrid-SR	0.776	0.594	0.282	0.071
RADDet (fine.) + RAD-SR	0.774	0.601	0.278	0.067
RADDet (fine.) + ADC-SR	0.794	0.620	0.304	0.079
RADDet (fine.) + Hybrid-SR	**0.802**	**0.631**	**0.312**	**0.085**

Table 6. Results of radar object detection using 2D YOLO head following RADDet [30]

Method	$AP_{0.1}$	$AP_{0.3}$	$AP_{0.5}$	$AP_{0.7}$
RADDet	0.801	0.730	0.530	0.202
RADDet (fix.) + RAD-SR	0.811	0.736	0.541	0.223
RADDet (fix.) + ADC-SR	0.805	0.738	0.539	0.227
RADDet (fix.) + Hybrid-SR	0.827	0.741	0.542	0.235
RADDet (fine.) + RAD-SR	0.832	0.760	0.577	0.252
RADDet (fine.) + ADC-SR	0.835	0.761	0.579	0.250
RADDet (fine.) + Hybrid-SR	**0.841**	**0.773**	**0.588**	**0.267**

plugging in our models without fine-tuning the detector already brings performance gains. Second, fine-tuning the detector using our pre-trained super-resolution models can further improve the performance by a large margin. This demonstrates that azimuth super-resolution models are suitable for detection tasks in autonomous driving.

2D YOLO head. We also evaluate the RADDet detector using different kinds of super-resolution models: RAD-SR, ADC-SR, and Hybrid-SR when producing the 2D bounding boxes in RA maps. The results are presented in Table 6. We can also observe that our models already bring performance gain without fine-tuning the detector. In addition, fine-tuning the detector using our trained models as data augmentations is able to bring additional performance gains. This indicates our models support both types of YOLO heads in the radar object detection task.

5.6. Visualization of RA maps

To qualitatively analyze the effectiveness of azimuth super-resolution of our models, we compare the RA maps produced by RAD-SR, ADC-SR, and Hybrid-SR with one baseline 3D SRGAN in Figure 6. This figure shows the outputs of RA maps when taking into the same low-resolution ADC signals or RAD maps as inputs. As we observe from Figure 6, there are some phenomena that can be summarized. First, 3D SRGAN struggles to reconstruct the high-resolution RA maps since it does not learn how to sharpen and separate the responses from low-resolution RAD maps. Second, RAD-SR employing RAD-Unet is able to roughly translate the low-resolution RAD maps to high-resolution RAD maps with image-to-image translation modules. Yet, though RAD-SR is able to produce the super-resolved RA

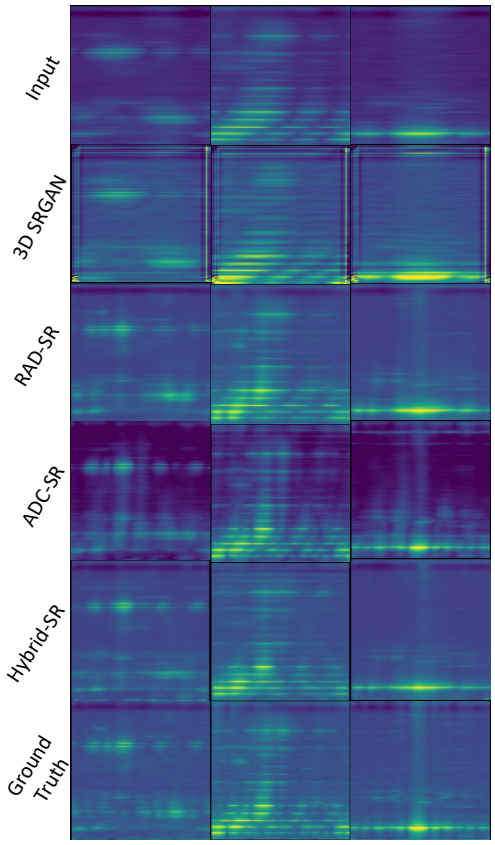

Figure 6. **Qualitative comparisons and analysis in reconstruction of Range-Azimuth (RA) maps.**

maps close to ground truths, it still fails to separate some regions of small responses. Third, ADC-SR and RAD-SR taking into ADC signals directly are able to sharpen the responses and have them well-matched to the ground truths. This demonstrates models taking into ADC signals are preferable for azimuth super-resolution.

6. Conclusions

To address the task of azimuth super-resolution in MIMO radar, a type of FMCW radar device, we propose an ADC super-resolution (ADC-SR) model for predicting unseen signals. Our ADC-SR model utilizes 50 times fewer parameters yet exhibits comparable performance than the baseline RAD super-resolution model, demonstrating the significance of leveraging ADC signals. Our Hybrid-SR combining our ADC-SR with RAD-SR improves the RAD-SR by a large margin on our collected dataset Pitt-Radar where ADC signals are available for evaluation. The experiments on plugging our models into the existing object detector also support the generalization of our models.

Acknowledgement: We thank Denso for the sponsorship and the data collection. We also thank Daiji Watanabe and Hironobu Akita from Denso for the project support.

References

[1] Navtech radar device technical description. https://navtechradar.com/clearway-technical-specifications. 3

[2] Karim Armanious, Sherif Abdulatif, Fady Aziz, Urs Schneider, and Bin Yang. An adversarial super-resolution remedy for radar design trade-offs. In *2019 27th European Signal Processing Conference (EUSIPCO)*, pages 1–5. IEEE, 2019. 2

[3] Dan Barnes, Matthew Gadd, Paul Murcutt, Paul Newman, and Ingmar Posner. The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6433–6438. IEEE, 2020. 3

[4] Mario Bijelic, Tobias Gruber, Fahim Mannan, Florian Kraus, Werner Ritter, Klaus Dietmayer, and Felix Heide. Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 11682–11692, 2020. 1

[5] Holger Caesar, Varun Bankiti, Alex H Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuscenes: A multimodal dataset for autonomous driving. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 11621–11631, 2020. 1, 2, 3

[6] Joel F Campbell, Bing Lin, Amin R Nehrir, F Wallace Harrison, and Michael D Obland. Super-resolution technique for cw lidar using fourier transform reordering and richardson–lucy deconvolution. *Optics letters*, 39(24):6981–6984, 2014. 2

[7] Cristóvão Cruz, Rakesh Mehta, Vladimir Katkovnik, and Karen O Egiazarian. Single image super-resolution based on wiener filter in similarity domain. *IEEE Trans. Image Process.*, 27(3):1376–1389, 2017. 2

[8] Semih Doğu, Mehmet Nuri Akıncı, Mehmet Çayören, and İbrahim Akduman. Truncated singular value decomposition for through-the-wall microwave imaging application. *IET Microwaves, Antennas & Propagation*, 14(4):260–267, 2020. 2

[9] Herbert Dropkin and Canh Ly. Superresolution for scanning antenna. In *Proceedings of the 1997 IEEE National Radar Conference*, pages 306–308. IEEE, 1997. 2

[10] Andrew Geiss and Joseph C Hardin. Radar super resolution using a deep convolutional neural network. *Journal of Atmospheric and Oceanic Technology*, 37(12):2197–2207, 2020. 2

[11] Texas Instruments. Short range radar reference design using awr1642. Technical report, Technical Report, 2017. 1, 4

[12] Giseop Kim, Yeong Sang Park, Younghun Cho, Jinyong Jeong, and Ayoung Kim. Mulran: Multimodal range dataset for urban place recognition. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6246–6253. IEEE, 2020. 3

[13] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017. 4, 5, 6, 7

[14] Peizhao Li, Pu Wang, Karl Berntorp, and Hongfu Liu. Exploiting temporal relations on radar perception for autonomous driving. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 17071–17080, 2022. 3

[15] Zhongyu Li, Shanchuan Li, Zhutian Liu, Haiguang Yang, Junjie Wu, and Jianyu Yang. Bistatic forward-looking sar mp-dpca method for space–time extension clutter suppression. *IEEE Transactions on Geoscience and Remote Sensing*, 58(9):6565–6579, 2020. 2

[16] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *Eur. Conf. Comput. Vis.*, pages 740–755. Springer, 2014. 6

[17] Mohammadreza Mostajabi, Ching Ming Wang, Darsh Ranjan, and Gilbert Hsyu. High-resolution radar dataset for semi-supervised learning of dynamic objects. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 100–101, 2020. 3

[18] Arthur Ouaknine, Alasdair Newson, Julien Rebut, Florence Tupin, and Patrick Perez. Carrada dataset: Camera and automotive radar with range-angle-doppler annotations. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 5068–5075. IEEE, 2021. 3

[19] Kun Qian, Shilin Zhu, Xinyu Zhang, and Li Erran Li. Robust multimodal vehicle detection in foggy weather using complementary lidar and radar signals. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 444–453, 2021. 3

[20] C Raju and T Sreenivasulu Reddy. Mst radar signal processing using iterative adaptive approach. *Geoscience Letters*, 5(1):1–10, 2018. 2

[21] Julien Rebut, Arthur Ouaknine, Waqas Malik, and Patrick Pérez. Raw high-definition radar for multi-task learning. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 17021–17030, 2022. 2, 3, 6

[22] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. 6, 7

[23] Ole Schumann, Markus Hahn, Nicolas Scheiner, Fabio Weishaupt, Julius F Tilly, Jürgen Dickmann, and Christian Wöhler. Radarscenes: A real-world radar point cloud data set for automotive applications. In *2021 IEEE 24th International Conference on Information Fusion (FUSION)*, pages 1–8. IEEE, 2021. 3

[24] Marcel Sheeny, Emanuele De Pellegrin, Saptarshi Mukherjee, Alireza Ahrabian, Sen Wang, and Andrew Wallace. Radiate: A radar dataset for automotive perception in bad weather. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–7. IEEE, 2021. 3

[25] Andrew G Stove. Linear fmcw radar techniques. In *IEE Proceedings F (Radar and Signal Processing)*, volume 139, pages 343–350. IET, 1992. 1, 3

[26] Shunqiao Sun, Athina P Petropulu, and H Vincent Poor. Mimo radar for advanced driver-assistance systems and autonomous driving: Advantages and challenges. *IEEE Signal Processing Magazine*, 37(4):98–117, 2020. 1, 2, 6

[27] Yizhou Wang, Zhongyu Jiang, Xiangyu Gao, Jenq-Neng Hwang, Guanbin Xing, and Hui Liu. Rodnet: Radar object detection using cross-modal supervision. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 504–513, 2021. 2

[28] Yizhou Wang, Zhongyu Jiang, Yudong Li, Jenq-Neng Hwang, Guanbin Xing, and Hui Liu. Rodnet: A real-time radar object detection network cross-supervised by camera-radar fused object 3d localization. *IEEE Journal of Selected Topics in Signal Processing*, 15(4):954–967, 2021. 2

[29] Yizhou Wang, Gaoang Wang, Hung-Min Hsu, Hui Liu, and Jenq-Neng Hwang. Rethinking of radar's role: A camera-radar dataset and systematic annotator via coordinate alignment. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 2815–2824, 2021. 3

[30] Ao Zhang, Farzan Erlik Nowruzi, and Robert Laganiere. Raddet: Range-azimuth-doppler based radar object detection for dynamic road users. In *2021 18th Conference on Robots and Vision (CRV)*, pages 95–102. IEEE, 2021. 2, 3, 4, 6, 7, 8

[31] Qiping Zhang, Yin Zhang, Yulin Huang, and Yongchao Zhang. Azimuth superresolution of forward-looking radar imaging which relies on linearized bregman. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7):2032–2043, 2019. 2

[32] Qiping Zhang, Yin Zhang, Deqing Mao, Yongchao Zhang, Yulin Huang, and Jianyu Yang. A bayesian super-resolution method for forward-looking scanning radar imaging based on split bregman. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 5135–5138. IEEE, 2018. 2

[33] Yongchao Zhang, Andreas Jakobsson, Yin Zhang, Yulin Huang, and Jianyu Yang. Wideband sparse reconstruction for scanning radar. *IEEE Transactions on Geoscience and Remote Sensing*, 56(10):6055–6068, 2018. 2

[34] Yongchao Zhang, Deqing Mao, Qian Zhang, Yin Zhang, Yulin Huang, and Jianyu Yang. Airborne forward-looking radar super-resolution imaging using iterative adaptive approach. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7):2044–2054, 2019. 2

[35] Yin Zhang, Qiping Zhang, Yongchao Zhang, Jifang Pei, Yulin Huang, and Jianyu Yang. Fast split bregman based deconvolution algorithm for airborne radar imaging. *Remote Sensing*, 12(11):1747, 2020. 2