This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043

IEEE Open Journal of Intelligent Transportation Systems

Received 10 April, 2023; revised 30 July, 2023; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version XX Month, XXXX.

Digital Object Identifier 10.1109/OJITS.2022.1234567

Automotive radar sub-sampling via object detection networks: Leveraging prior signal information

Madhumitha Sakthi, *Student Member, IEEE*, Marius Arvinte, and Haris Vikalo, *Senior Member, IEEE*

[1]Department of Electrical and Computer Engineering, The University of Texas at Austin

CORRESPONDING AUTHOR: Madhumitha Sakthi (e-mail: madhumithasakthi.iyer@utexas.edu).

ABSTRACT In recent years, automotive radar has attracted considerable attention due to the growing interest in autonomous driving technologies. Acquiring situational awareness using multimodal data collected at high sampling rates by various sensing devices including cameras, LiDAR, and radar requires considerable power, memory and compute resources which are often limited at an edge device. In this paper, we present a novel adaptive radar sub-sampling algorithm designed to identify regions that require more detailed/accurate reconstruction based on the information about prior environmental conditions, enabling near-optimal performance at considerably lower effective sampling rates. Designed to robustly perform under variable weather conditions, the algorithm was shown on the Oxford radar dataset to achieve accurate scene reconstruction utilizing only 10% of the collected samples in good weather. In the case of the RADIATE dataset acquired during extreme weather conditions (snow, fog), only 20% of the samples were sufficient to enable robust scene reconstruction. A further modification of the algorithm incorporates object motion to enable reliable identification of regions that require attention. This includes monitoring possible future occlusions caused by the objects detected in the present frame. Finally, we train a YOLO network on the RADIATE dataset to perform object detection, obtaining 6.6% AP50 improvement over the baseline Faster R-CNN network.

INDEX TERMS Compressive sensing, automotive radar, sub-sampling, Faster R-CNN, YOLO, object detection, measurement matrix, signal acquisition

I. Introduction

MANY sensing technologies commonly used by autonomous vehicles, including cameras and LiDAR, are generally capable of providing situational awareness but struggle under extreme weather conditions (e.g., heavy rain, fog and snow) [1]. Automotive radar is a robust all-weather choice that can provide reliable information about the locations of objects in the environment. For instance, broadly used Frequency-Modulated Continuous Wave (FMCW) radar operates in the wide 30-300 GHz band to provide highly accurate range data [2]. Prior research includes using radar data to extract point clouds which are subsequently used for object detection [3]–[6]. Applications of radar include the prediction of vehicle behavior and traffic flow state detection [7], [8], among others; the methods relying solely on radar data have been shown to enable robust object detection, making radar an attractive choice for acquiring situational awareness in extreme weather conditions [2], [9]–[14].

To achieve desired target accuracy, radars typically collect data at very high rates. For instance, the aforementioned FMCW radar [15] operating at 76-77 GHz generates data at about 16 Gbps. The analog-to-digital converter (ADC) thus needs to function at an extremely high rate to capture the reflected signal and forward it to the DSP for processing [16]. The time-domain signal is then converted to the frequency domain, resulting in a Range-Azimuth-Doppler tensor that needs to be further processed to identify the location of the objects. This information may need to be communicated (vehicle-to-vehicle and/or vehicle-to-everything) to help multiple road actors acquire situational awareness. To relax the computational and energy requirements on the data processing pipeline, and to transmit data within latency constraints, it would be beneficial to reduce the data rate without compromising the quality of the acquired situational awareness. To this end, previous studies have explored the

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043

use of compressed sensing (CS) methods to enable trade-offs between sampling rate and acquisition quality [17].

In this paper, we present a method for the adaptive acquisition of range-azimuth radar data via compressed sensing aided by the results of object detection from previous frames. In particular, the proposed algorithm splits each radar frame into uniform-sized blocks and relies on object detection to determine the blocks that require more attention. Given a total sampling budget, the algorithm solves an aptly formulated linear program to adaptively allocate a higher sampling rate to the blocks that require more attention, and a lower sampling rate to the remaining ones.

The proposed framework assumes the general availability of both camera and radar data and adapts the sampling strategy to varied weather conditions. In good weather, the sub-sampling algorithm uses the previous frame's data (images, radar) to decide the importance of regions in the incoming radar data; the object detection algorithm runs on the images and identifies bounding boxes and objects' classes. These bounding boxes are used to determine the initial set of important regions. If an object is either missed by the image-based object detection network or is in the blind spot of a camera, previous radar data are used to assist in identifying regions of interest. After converting the results of the object detection step from image to radar coordinates, a linear program (LP) is formulated and solved to help identify sampling rates for each radar block. In bad/extreme weather conditions, the algorithm does not rely on image sensors to determine regions of interest but rather uses as prior information the positions of tracked objects predicted via a Kalman filter [18]. While previous work [19] has shown accurate performance in motion prediction tasks using camera and Lidar, these sensors are prone to failure during extreme weather conditions. This motivates the development of a robust acquisition pipeline utilizing radar data and its application to downstream tasks such as perception and motion prediction.

The main contributions of the paper are as follows:

- A novel adaptive algorithm is designed to facilitate radar data acquisition by relying on both images and radar in good weather conditions, and radar only in adverse weather conditions.
- The proposed radar-only data acquisition algorithm is designed to select regions that are more likely to contain objects, ultimately enabling accurate performance at considerably lower sampling rates. The method relies on object motion prediction to facilitate the selection of regions that require more attention.
- The important region selection process is designed to account for possible occlusions in future frames by the objects detected in the current frame.
- Experimental verification via the You Only Look Once (YOLO) [20] object detection network trained on the RADIATE [15] dataset demonstrates a significant improvement in terms of the AP50 metric over a baseline Faster R-CNN network (63.8% vs. 57.2% AP50).
- To the best of our knowledge, this is the first work to investigate spatially adaptive radar sub-sampling on the automotive FMCW real-world radar data.

The remainder of the paper is organized as follows. Section II provides an overview of the existing efficient acquisition methods followed by the description of object detection networks trained on images and radar point cloud and radar-based object detection. The proposed algorithms are presented in Section III, along with the details of the compressed sensing procedure. Section IV presents experimental results on the Oxford robocar [21] and RADIATE [15] datasets. The paper is concluded in Section V. The preliminary results of this work were reported in [22], [23].

II. Related work

A. Compressed sensing for radar systems

The compressed sensing (CS) techniques enable acquisition of data at sub-Nyquist sampling rates by striking a trade-off between sampling rate and reconstruction quality. Applying CS ideas, the work in [24] presents reliable reconstruction of Synthetic Aperture Radar (SAR) data using 70% of the points in the full dataset. Similarly, in [25], CS is utilized for noise radar data reconstruction using only 30% of the reference signal. In [26], the authors present an efficient CS-based reconstruction of a frequency-modulated continuous wave (FMCW) radar using 40% of the original samples. In [27], the authors analyzed various CS reconstruction algorithms including Orthogonal Matching Pursuit (OMP) and Basis Pursuit De-noising (BPDN), and have shown that, in automotive settings, OMP exhibits superior performance. Our proposed framework relies on the Basis Pursuit (BP) algorithm for signal reconstruction since BP, though more computationally expensive, generally requires fewer measurements than OMP [28].

Adaptive CS is a broadly used technique for increasing sampling rate in the regions deemed important [29]–[31]. The scheme in [29] utilizes previously received pulse interval and relies on the constant false alarm rate (CFAR) to decide how to subsample pulsed radar data. In [32], the authors introduced a compressed sensing-based CFAR algorithm for improving the detection performance without reconstructing the radar data; in contrast, the current paper aims at reconstructing the radar data that can be utilized for further downstream tasks including object detection and the segmentation of the range-azimuth radar data using neural networks. In [30], an adaptive CS algorithm was used to improve target tracking performance in static settings. The authors of [31] aim at optimizing the measurement matrix in settings where only targets are moving, improving performance at the cost of increased computational complexity. In contrast, this paper investigates the application of adaptive CS for radar acquisition in settings where the tracked objects and the vehicle deploying the sensors are both potentially moving. A

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043

further distinctive feature of our work is that we allocate to the important regions of the environment a larger sampling budget while maintaining the overall sampling budget and reconstruction complexity. In another related work [33], the acquisition of LiDAR data is guided using the Region-of-Interest information determined based on the results of image segmentation; in contrast, our work helps guide radar data acquisition utilizing the results of 2-D object detection.

The ADC's sampling frequency and resolution play a major role in the acquisition of the intermediate frequency (IF) signal following the signal mixing stage. In order to limit the rate of the received radar data, the authors of [34] used a bit-limited MIMO radar with an additional analog filter to form a global hybrid analog-digital system. In [35], high-resolution ADCs were replaced by one-bit ADCs using time-varying thresholds. In order to estimate the angle and Doppler frequency from one-bit sampled data, [36] uses a maximum-likelihood-based method. In contrast, our method relies on DNN output to assist in efficient spatial sampling, thereby focusing on the regions of interest while limiting the overall bit budget during acquisition without needing analog filters or threshold-modified ADC.

When compressively sampling a signal, the choice of a measurement matrix is critical in ensuring the signal's reconstruction. In addition to often used Gaussian matrices, binary measurement matrices received considerable attention due to their hardware advantages. In [37], the authors proposed a Binary Permuted Block Diagonal (BPBD) measurement matrix with equally-sized diagonal blocks randomly permuted along the columns. There, BPBD was compared with the scrambled Fourier and Partial Noiselet alternatives, among others, showing comparable reconstruction performance on images. In related work, [38] proposed a CS reconstruction approach that relies on an extended smoothed-projected Landweber algorithm. In [39], binary random measurement matrices were used for CS of ECG signal, while in [40] a binary block diagonal matrix without permutation was introduced as a deterministic measurement matrix for the compression and recovery of electrocardiogram (ECG) and Electromyography signals. In [41], the authors applied a binary measurement matrix to the problem of investigating soils or stone walls, showing successful reconstruction with reduced sampling rates, while [42] used a binary measurement matrix to facilitate compressive measurements in DNA microarrays. Evidently, the binary measurement matrix has been successful in a variety of CS applications. In this paper, we rely on binary permuted diagonal (BPD) matrices for automotive radar data acquisition and report the resulting object detection performance in terms of the AP and AP50 metrics.

B. Object detection using radar and images

Previous studies demonstrated that combining radar data with images helps improve object detection performance. In [43], a spatial attention fusion method was proposed; the method was embedded at the feature-extracting stage to effectively leverage both radar and camera features. In [44], the authors deployed a Faster R-CNN object detection network, replacing the selective search-based region proposal algorithm with the region proposals generated using radar data points. That work experimented on the NuScenes [45] dataset, performing selective search to show AP improvement from 41.8% to 43.0%. The authors in [4] report that using the combination of radar and images enables more accurate 3-D object detection performance than a scheme that relies on LiDAR and images. In [3], the authors used a center point detection network and proposed a frustum-based method to facilitate 3-D object detection, showing improved occluded object detection. In [43], radar and image features were combined via spatial attention to improve the object detection performance while using the FCOS [46] object detection network pipeline.

C. Object detection using radar-only

This line of research focuses on generating object detection networks from radar data without relying on any other sensors. To this end, several radar datasets with object annotations have been collected and open-sourced [2], [15]. In [2], the authors proposed a ROD-Net network for radar object detection that resulted in 86% average precision; they also released CRUW, a dataset containing object annotations on RF images. The method in [9] combined particle filter-based tracking and object detection to enable radar-based object identification, demonstrating accurate performance while alternating between tracking and detection to reduce the computational load induced by real-time processing. In [12], a channel boosting feature ensemble method was proposed and validated on the RADIATE dataset; the method uses a transformer neural network for direct object detection from radar frames. However, this approach requires bulky backbone networks such as resnet-50 or resnet-101 to process RGB, LUV and LAB radar frames individually, rendering its practical applications challenging.

III. Method

The diverse environments in day-to-day settings, with weather conditions ranging from sunny to extreme, present distinct challenges and require robust algorithms for radar signal acquisition. Fig. 1 shows a block diagram of the possible inputs and the subsequent signal processing pipeline. Regardless of the inputs, in the considered setting an object detection network is used and the radar data is acquired via compressed sensing; depending on the input signal, the linear programming problem that assists in identifying sampling rates is formulated differently. The details of each pipeline are elaborated upon in the following sections.

A. The Compressed Sensing Framework

The proposed method splits the measured frame into multiple blocks and probes each block with a predetermined number

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043

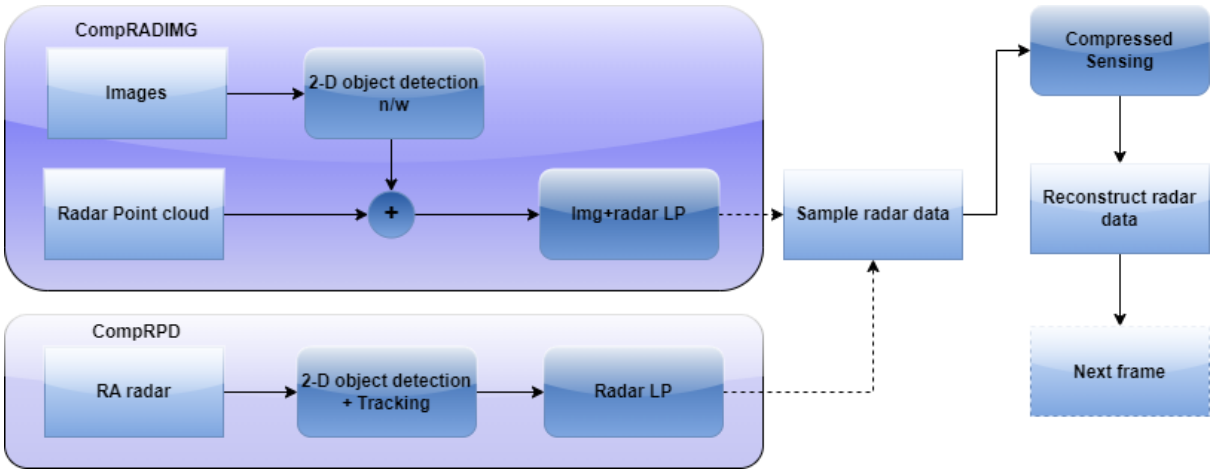


FIGURE 1. An illustration of the radar sub-sampling scheme for processing radar data in either good or bad weather conditions. The CompRADIMG block is used in good weather conditions to compressively sample radar signal using both prior image and radar data. The CompRPD block is used for compressive sampling utilizing only the previous radar data. Following the sampling process, the desired signal is reconstructed using the Basis Pursuit (BP) algorithm.

of measurements (e.g., m measurements of an n-dimensional signal are collected from each block). This procedure is predicated on two assumptions: the signal is sparse in some domain and the measurement matrix satisfies the restricted isometry property [47]. Specifically, we assume that the data is sparse in the Discrete Cosine Transform (DCT) domain. Given the original signal $x \in R^n$, each block is probed with a random BPD measurement matrix $\phi \in R^{m \times n}$ resulting in measurements $y \in R^m$. Formally, the measurements y are of the form

$$y = \phi x. \tag{1}$$

Given the measurements, the desired signal x is recovered by running the Basis Pursuit (BP) algorithm [48], [49]. In particular, we use the publicly available L_1-MAGIC toolbox [50] to solve the constrained optimization problem

$$\begin{aligned} &\min_{x} \ \|\theta x\|_1 \\ &\text{s.}t. \ \ \phi x = y, \end{aligned} \tag{2}$$

where θ denotes the DCT transformation matrix [51].

The Gaussian measurement matrix, typically used to acquire a linear combination of the signals, requires floating-point multiplications. The authors in [37] successfully replaced the Gaussian measurement matrix with a Binary Permuted Block Diagonal (BPBD) measurement matrix without major performance deterioration. In [23], the same idea was followed to introduce Binary Permuted Block (BPD) measurement matrices that help avoid power-consuming complex multipliers and instead rely on hardware-efficient elements including switches and selectors as in [38]. The performance induced by different measurement matrices is characterized in terms of AP and AP50, the task-relevant metrics, as opposed to the PSNR typically shown in prior work.

The baselines used in the benchmarking experiments include: (1) the sampling framework where the samples/bits are uniformly distributed across the frame being sampled, and no prior frame information is used [25]; and (2) the CFAR algorithm which determines important regions and utilizes this information to dynamically allocate sampling budget across the frame [29].

B. CompRADIMG: Sub-sampling using previous images and radar data

In [23], the images from the camera are first processed by the Faster R-CNN network to identify the bounding boxes and classes of the objects. The network's output is then converted from the camera coordinates to the radar azimuth blocks. The sampling rate used during the subsequent acquisition of radar data is dynamically determined by running a linear program specified later in this section. This entire procedure, lasting in total ~ 0.25s, is repeated for the upcoming frames. However, a radar compressive sampling strategy that relies only on prior image data, such as the one in [23], may miss an object if the object appears in the blind spot of the camera or, more generally, if the object detection network fails to detect it. This could lead to serious consequences in downstream applications that heavily rely on radar data for depth estimation, path planning and other objectives. Aiming to address such scenarios and generally improve the performance of a compressive sensing scheme, we introduce the CompRADIMG (Compression using RADar and IMaGes) algorithm which identifies regions with the objects of interest by combining information present in the previous image and radar data. To this end, CompRADIMG converts the bounding boxes of the objects in the image coordinates to the azimuth in the radar frames. The azimuth regions containing objects, identified by CFAR, are also provided to the linear program that determines acquisition rates for the blocks in the present radar frame. Specifically, a radar frame covering the range and azimuth of the scene is uniformly partitioned into blocks where each block covers a certain range and azimuth; we would like to determine the acquisition rate

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043

for each of these blocks. Note that an area covered by an object such as a pedestrian or bicycle is typically smaller than that covered by a car or a bus (due to the smaller area of reflection). To enable accurate reconstruction of the areas with such small objects, the algorithm assigns higher importance (effectively, the largest sampling rate) to those areas, followed by the motorized vehicles and, lastly, other objects (roads, walkways) in the scene. To this end, the blocks are organized in regions; the entire radar frame is divided into four regions, as explained next.

We divide the range of the radar frame (typically exceeding 100m) in two parts: (1) the first part ("first range"), which extends from the ego vehicle to the half-range, and (2) the second part ("second range"), accounting for the rest. Now, the azimuth in the first range of the radar is split into three regions. Let a_{11} denote the total number of blocks with a small object such as pedestrian or bicycle within the first range (as defined before). Similarly, let a_{12} and a_{13} denote the total number of blocks with cars or other objects within the first range, respectively. The total number of blocks in the second range across the entire azimuth is denoted by a_2. Then the first region is defined as the union of (generally non-adjacent) blocks counted towards a_{11}; the remaining regions are defined likewise only in regards to the blocks associated with a_{12}, a_{13}, and a_2.[1] Let us define

$$f(x) = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_2x_4, \tag{3}$$

where x_1, x_2, x_3, x_4 denote the relative acquisition rates (i.e., fractions of the full acquisition rate) for the four regions we defined, respectively. Then the acquisition rate for these regions is determined by solving

$$\begin{aligned} \max_{x_1,x_2,x_3,x_4 \geq 0} \; & f(x) \\ s.t. \; & x_1 = \alpha x_3, \; x_2 = \beta x_3 \\ & f(x) \leq S, \; x_{il} \leq x_i \leq x_{iu} \quad \text{for } i = 1,2,3,4. \end{aligned}$$

The constraint $f(x) \leq S$ limits the total number of samples. Since x_1 is the acquisition rate for the blocks containing small (important) objects, it is allowed to have the highest sampling rate (as facilitated by $x_1 = \alpha x_3$, $\alpha > 1$). Similarly, the constraint $x_2 = \beta x_3$ ($\beta > 1$) ensures that the region with a large object (e.g., a motorized vehicle) is allocated a higher sampling rate than the road/walkway. The lower constraints on the sampling rates enforce that a sufficient number of samples is collected to enable reconstruction.

The pseudo-code for the described procedure is given as Algorithm 1. The function f_{NN} is the pre-trained object detection network that predicts bounding boxes in the images. The function f_{CFAR} [52] determines the important blocks across every row in the radar frame $\hat{X}_t$. The azimuth of an object's bounding box is obtained from its image coordinates according to $f_A(B) = (centre_x/X_{width})(\theta_{max} - \theta_{min})$, where $centre_x$ denotes the x-coordinate of the detected object bounding box's center, θ_{min} and θ_{max} are the minimum and maximum field of view of the camera relative to the position of the camera in the bird's eye view frame, respectively, and X_{width} is the total width of the image. After determining the sampling/acquisition rates, the frame is compressively sampled and reconstructed using the compressed sensing algorithm f_{CS}.

[1]Note: to allow for additional flexibility when forming the regions, if the CFAR algorithm declares that an object originally categorized as contributing to a_{12} or a_{13} is exceptionally important, the corresponding block is reclassified so that it instead contributes to a_{11}.

Algorithm 1 CompRADIMG Algorithm

Input: radar frame $X \in \mathbb{R}^{m \times n}$ and image frame I
Output: Reconstructed radar frame $\hat{X}$
At $t = 1$, Initialize first radar frame $\hat{X}_t$
for $t = 2, \ldots, T$ **do**

1) Determine object bounding boxes $B \leftarrow f_{NN}(I_{t-1})$
2) Determine important image blocks $b \leftarrow f_{CFAR}(\hat{X}_{t-1})$
3) Determine important azimuth $A \leftarrow f_A(B)$
4) Determine sampling rate for t-th frame X_t using $x_1, x_2, x_3, x_4 \leftarrow f_{LP}(A, b)$
5) Output compressed and reconstructed $\hat{X}_t \leftarrow f_{CS}(x_1, x_2, x_3, x_4, X_t)$

end for

C. CompRPD: Radar sub-sampling using previous radar

In adverse weather conditions, a sub-sampling algorithm may only rely on previous radar data to identify important regions in the present frame. Extending the work in [22], we present CompRPD (Compression of Radar using Predictions and Detections), an algorithm for identifying important regions in such settings. In [22], the radar data was split into equal-sized blocks and, based on the detection result, locations of the objects of interest were identified. The algorithm presented in the current paper, however, additionally relies on the locations of the objects predicted based on their prior location – the information available due to running the algorithm over multiple frames.

In particular, we use a Kalman filter to predict the location of an object in the present frame based on the analysis of the previous frame. The Kalman filter is initialized using the information about bounding boxes of the object detected in the first (anchor) frame, sampled at a higher sampling rate and processed by the object detection network. In the next frame, bounding boxes identified by the object detection network are compared in terms of the Intersection over Union (IOU) with the corresponding bounding boxes predicted by the Kalman filter; the result of this comparison is then utilized to perform the measurement update step in the Kalman filter recursion. The predicted bounding box location is used to guide the sampling rate decision unless the confidence of the prediction, quantified by means of the aforementioned IOUs or the lack of identified matches, exceeds a predetermined hyperparameter referred to as the

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043

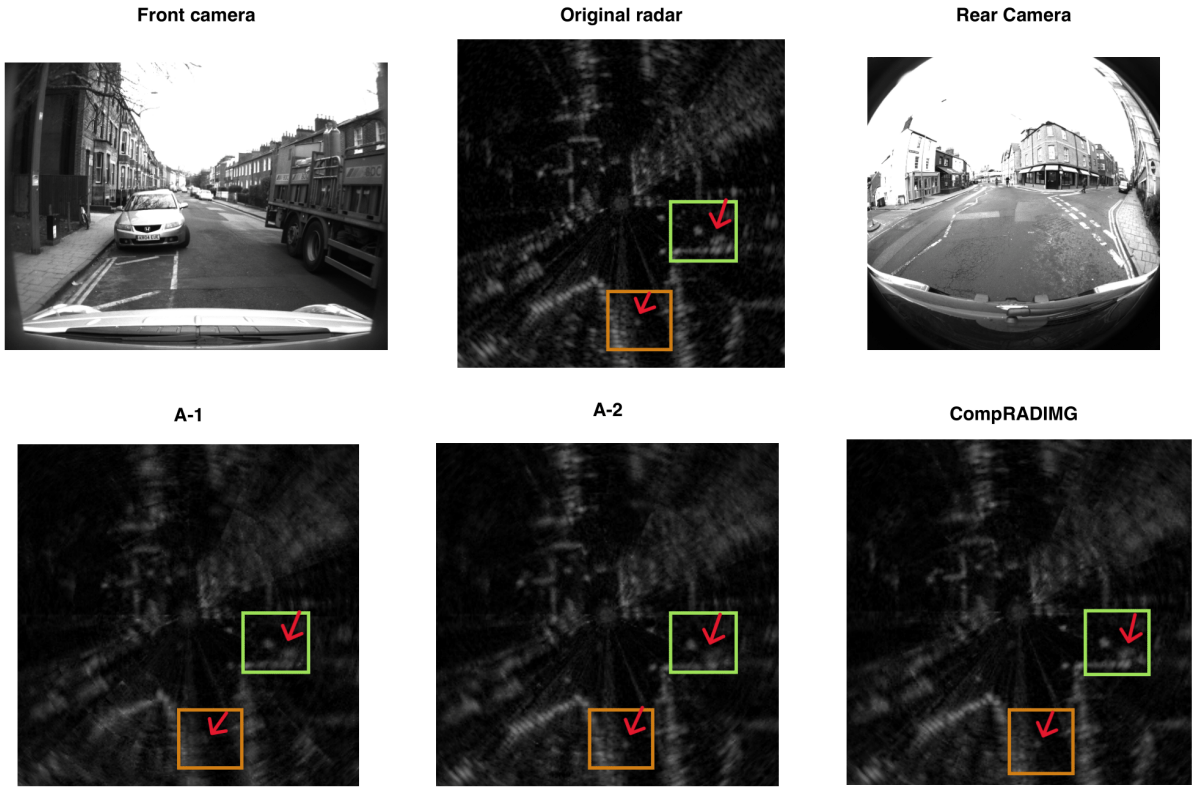


FIGURE 2. The original radar and the corresponding image sequence along with the results of reconstruction by the CompRADIMG algorithm (scene 2, frame 9, Oxford dataset). The green box indicates a car to the rear right missed by the baseline A-2 but detected/reconstructed by CompRADIMG. The orange box indicates the bike detected by CompRADIMG but missed by the baseline A-1. *Best viewed as a digital copy by zooming in.*

maximum age; in the latter case, the sampling rate decision is made based directly on the object's bounding box detected in the current frame. Overall, the signal acquisition rate is determined based on either the object's current position or its predicted future position. In either case, the object's coordinates are first converted into radar polar coordinates, followed by specifying the blocks of interest. For this, the algorithm relies on an inverted T-shaped sampling pattern which takes into account the specifics of the Cartesian to polar block conversion to ensure that the space partition into blocks is such that an adequate amount of information is acquired from the surroundings of the autonomous vehicle.

Let us define

$$f(x) = I \cdot x_1 + O \cdot x_2, \tag{4}$$

where I is the total number of important blocks, O is the total number of other blocks, and x_1 and x_2 denote the relative acquisition rates (i.e., fractions of the full acquisition rate) for important and other blocks, respectively. Then, the linear program is used to dynamically determine the sampling rate for each block in the current radar frame can be stated as

$$\begin{aligned} &\max_{x_1, x_2 \geq 0} \ f(x) \\ &\quad s.t. \ x_1 \geq \alpha x_2 \\ &\qquad f(x) \leq S, \ x_{il} \leq x_i \leq x_{iu} \quad \text{for } i = 1, 2. \end{aligned}$$

The constraint $f(x) \leq S$ limits the number of samples, e.g., to $10 - 30\%$ of the full acquisition rate. For the constraint $x_1 \geq \alpha x_2$, we typically set $\alpha = 1.1$; this promotes the sampling rate in the important regions to exceed the sampling rate in the other regions. The upper bound on x_1 corresponds to the sampling rates achieving performance in terms of AP/AP50 metrics matching that of the full acquisition; the lower bounds are reflective of the total sampling budget given to the entire frame so that the reconstruction quality of the important blocks at least matches that achieved when the entire area is sampled uniformly. As for x_2, the upper bounds are indicative of the total sampling budget while the lower bounds help ensure that a sufficient number of samples is collected to support reconstruction. Once the sampling rates are determined by solving the LP, they are utilized in the acquisition of the subsequent radar frames. When after processing a prespecified number of frames another anchor frame is encountered, the entire area is sampled at a full acquisition rate and the procedure starts anew.

CompRPD is formalized as Algorithm 2. The function f_{NN} is the pre-trained object detection network that predicts object bounding boxes based on the input radar Cartesian images. The $fullRate$ indicates radar frames acquired by sampling uniformly at the maximum acquisition/sampling rate. The function f_{CtoP} converts the bounding boxes from Cartesian to polar coordinates and specifies the central block.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043

Algorithm 2 CompRPD Algorithm

Input: Radar frame $X \in \mathbb{R}^{m \times n}$
Output: Reconstructed radar frame $\hat{X}_t$
For $t = 1$, initialize first radar frame $\hat{X}_t$
for $t = 2, \ldots, T$ **do**

1) Determine object bounding boxes $B \leftarrow f_{NN}(X_{t-1})$
2) **if** $t-1$ in $fullRate$ **then**
3) Initialize filter $K_{1,\ldots,len(B)}$ with $NNPred_{BB}$
4) $Final_{BB} \leftarrow B$
5) **else**
6) Determine Kalman predictions $KPred_{BB} \leftarrow Pred(K_{t-1})$
7) $IOU = f_{IOU}(KPred_{BB}, NNPred_{BB})$
8) **if** $IOU > 0$ **then**
9) Refine K using $NNPred_{BB}$
10) Update filter K using $NNPred_{BB}$
11) Update K_{age} by 1
12) **if** $K_{age} > MinAge$
13) $FinalBB$ = $KPred_{BB}$
14) **else**
15) $FinalBB$ = $NNPred_{BB}$
16) **end if**
17) **end if**
18) **end if**
19) Determine polar coordinates $G \leftarrow f_{CtoP}(Final_{BB})$
20) Determine sampling blocks $I \leftarrow f_{Blocks}(G)$
21) Set all other blocks to O
22) Determine sampling rate for t-th frame X_t using $x_1, x_2 \leftarrow f_{LP}(I, O)$
23) Output compressed and reconstructed $\hat{X}_t \leftarrow f_{CS}(x_1, x_2, X_t)$

end for

Depending on the size and the location of the object, the surrounding blocks are marked as important using the function f_{Blocks}. The function f_{LP} denotes the aforementioned linear program and is used to decide sampling rates. Finally, f_{CS} is the function performing compressive sampling and reconstruction of the current frame.

IV. Experimental Results

A. Data and baselines

The data used in the experiments include the Oxford radar robocar dataset [21] and the RADIATE dataset [15]. The Oxford dataset consists of Navtech CTS350-X Millimeter-Wave FMCW radar data along with the camera data from Point Grey Bumblebee XB3 rear camera, Point Grey Grasshopper2 stereo front camera and other sensor data. The data is collected on 280km of roads in Oxford, UK, with the front camera capturing at 16Hz frames per second (FPS), the rear camera capturing at 17Hz, and the radar capturing at 4Hz. To our knowledge, this is the only publicly available range-azimuth radar dataset accompanied by both front and rear camera data. Since the object annotations are not provided, we manually annotated a subset of it and tested the proposed algorithm on three random scenes with about 11 frames each.

The RADIATE dataset [15] was collected in extreme weather conditions including snow, fog, and rain. The dataset contains radar, LiDAR, camera and GPS data. The radar data was collected using Navtech CTS350-X with a 360° Horizontal Field of View (HFoV) and 100-meter range at 4Hz. This resulted in range-azimuth images of size 400x576 where the rows represent the angle and the column represents the range. The authors of [15] released 300 hours of annotated radar data, for both good and bad weather conditions, with annotation on the Cartesian radar images. Similar to [15], in our experiments we classify objects as vehicles or background; the vehicle classes include car, bus, bicycle, truck, van and motorbike. To test the proposed algorithm, we selected from the RADIATE dataset 40 frames for each snow, fog, motorway, city and night condition.

For the first benchmarking experiment, the radar frame is uniformly sampled at a rate determined without relying on any prior knowledge about the scene via compressed sensing (CS) [25] and the signal is recovered using BP. In the second experiment, the results of Cell Averaging-CFAR (CA-CFAR) [29] were used to identify radar blocks with objects of interest, and the sampling rate for those blocks was adaptively increased while maintaining the overall 10% sampling budget. This method is extended to Greatest Of-CFAR (GO-CFAR) [53] in place of CA-CFAR.

B. The measurement matrix for CS

To set up the stage for compressive radar acquisition, one needs to identify a suitable measurement matrix for rotating FMCW radar data. In [23], the BPD matrix for automotive radar data acquisition was proposed and the reconstruction quality was measured in terms of PSNR; BPD was found to achieve PSNR similar to that provided by the Gaussian and BPBD matrices. However, since radar object annotations are provided for the RADIATE dataset, we are able to test the performance induced by a specific measurement matrix in terms of the AP and AP50 metrics that are more relevant to the object detection task than PSNR. We reconstructed 5 scenes from the data with 40 frames each utilizing a standard CS algorithm, uniformly sampling across the frame at 10%, 20% and 30% relative sampling rates using three measurement matrices and Faster R-CNN and YOLOv5l networks. The results are shown in Table 1. The BPD measurement matrix outperforms the Gaussian and BPBD matrices in terms of the AP and AP50 metrics; the performance improvement is more significant at lower sampling rates than at higher sampling rates. Furthermore, YOLOv5 network outperformed Faster R-CNN at 30% acquisition rate while in the other cases, Faster R-CNN performs better. In general, evaluating performance in terms of task-relevant metrics such as AP or AP50 in addition to PSNR is ben-

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043

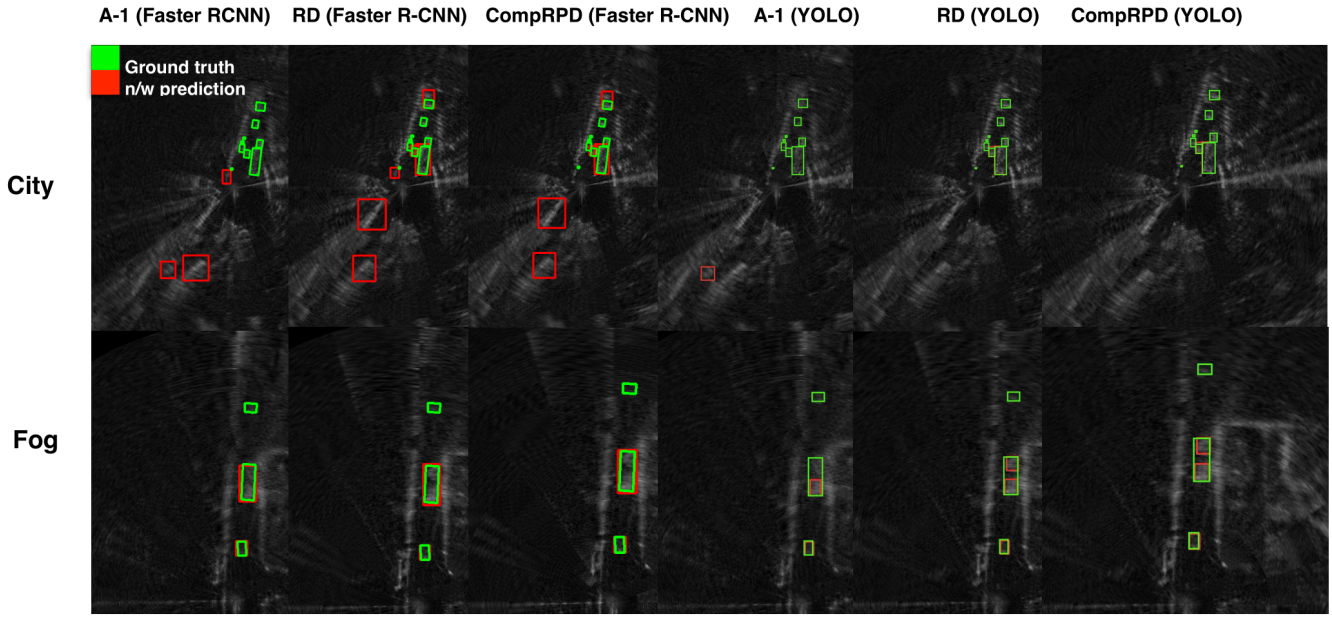


FIGURE 3. **The radar reconstruction and object detection results on two frames from city and fog scenes via the radar-detect (RD) [22] and CompRPD algorithms. The object annotation GT is in green and the network predictions are in red.**

eficial to identifying the best suited measurement matrix for compressive automotive radar sampling.

TABLE 1. **The AP, AP50 and PSNR results on the Oxford radar dataset [21] reconstructed using a standard CS algorithm. The objects were detected using both Faster R-CNN and YOLOv5 models.**

Sampling rate	Method	Gaussian	BPBD	BPD
10%	Faster R-CNN	5.5/2.2	5.6/2.0	**10.0/3.6**
	YOLOv5	3.1/1.4	3.7/1.6	7.1 / 2.4
	PSNR	25.39	25.65	25.4
20%	Faster R-CNN	31.0/12.3	32.3/12.6	**37.7/15.0**
	YOLOv5	27.6/11.2	29.3/11.8	36.3/14.2
	PSNR	26.58	26.68	26.59
30%	Faster R-CNN	47.8/19.3	48.2/19.8	49.2/20.3
	YOLOv5	50.5/20.8	50.4/20.9	**54.3/22.0**
	PSNR	27.66	27.71	27.66

C. CompRADIMG: Subsampling using image and radar

The results of testing CompRADIMG on the Oxford [21] dataset are reported in Table 2. We randomly selected three scenes, each with 11 frames. The frames were compressively sampled using only 10% of the original samples. The data, captured at 4Hz frame rate, has a range resolution 4.38cm and azimuth resolution 0.9°, with the total range of 163m. The rear camera has 180° horizontal field-of-view (HFoV) while the front camera's HFoV is 66°, with a blind spot of 57° on either side of the vehicle. The baseline algorithms include the algorithm in [25] (Labeled as A-1), CA-CFAR [29], GO-CFAR [29], [53] and the algorithm in [23] (Labeled A-2). The CFAR algorithms are applied on each row (range) to identify additional important bins with 300 training cells, 50 guard cells and a 0.001 false alarm rate.

In the first scene, the ego vehicle was driven on a straight road with pedestrians on either side of the walkway. Additionally, a van was following the ego vehicle, a bus was coming towards it, and a pedestrian was crossing the road in front of the ego vehicle. In frame 2, the person to the right was missed by the A-2 algorithm [23]. However, the person was detected by the CompRADIMG algorithm because the location the person was in was deemed to be important by the CFAR algorithm and hence was sampled at a higher rate. However, in frame 6, the person was missed by CompRADIMG, possibly due to the sampling budget constraints.

In the second scene, the car was passing through an intersection with a parked car to the front left and a truck passing by to the right. At the rear of the vehicle, there were multiple bikes at a distance and a pedestrian crossing on the rear left. In a few frames, the cars parked to the right of the vehicle appear in the camera's blind spot. In Figure 2, we show frame 9's radar and camera images. In this scene, the car parked to the right side of the vehicle was not reconstructed by the A-2 algorithm since it was located in the camera's blind spot. In contrast, the CompRADIMG algorithm successfully detected/reconstructed the car. The same car is also successfully reconstructed by the CompRADIMG algorithm in frames 6 and 8-9 of scene 2.

In the third scene, the vehicle was passing a crowded intersection with a bus to the front and multiple buses to the rear left. There were several pedestrians on either side of the road as well as a few bicycles. The CompRADIMG algorithm reconstructed the bike to the right in frames 7 and 10, which was missed by the A-2 algorithm. However, in frame 11 the person to the front left was missed by

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043.

TABLE 2. A comparison between CompRADIMG and the competing methods including A-1 [25], CA-CFAR [29], GO-CFAR [29], [53], and A-2 [23]. The presence of an object is indicated as 'yes'; if the object is faint or absent, the indicator is 'no'. Short notation: F-Front, Re-Rear, FL-Front Left.

Scene	Frame	Object	A-1 [25]	CA-CFAR [29]	GO-CFAR	A-2 [23]	CompRADIMG
1	2	Person (R)	no	yes	no	no	yes
	3	Person(FL)	no	no	no	yes	yes
	6	Person (R)	no	yes	yes	yes	no
	9	Person (R)	no	no	no	yes	yes
2	2-5,8,9,11	Bike (Re)	no	no	no	yes	yes
	6,8	Car (R)	no	no	no	no	yes
	9	Car (R)	no	yes	no	no	yes
	7	Car (R)	no	no	no	yes	no
3	2	Person(R)	no	yes	no	no	yes
	6	Person(R)	no	yes	yes	no	yes
	4	Car (R)	no	no	no	yes	yes
	7	Bike (R)	no	no	no	no	yes
	7	Person (R)	no	no	no	yes	yes
	10	Bike (R)	no	yes	yes	no	yes
	11	Person(FL)	no	no	no	yes	no

the CompRADIMG algorithm possibly due to an overly restrictive sampling budget.

D. CompRPD: Radar sub-sampling using previous radar

We tested the CompRPD algorithm on the city, motorway, snow, fog and night scenes of the radiate dataset [15]. The 1st and the 21st frames were set as anchor frames. We report results for three relative sub-sampling rates, 10%, 20% and 30%, where the lower bound on x_1 was set to 0.1, 0.2, and 0.3, respectively, while the upper bound was set to 0.55. These lower bounds ensure that the reconstruction quality of the important blocks is no worse than in the case the frame was uniformly sampled at the same overall sampling budget. The upper bounds on x_2 were set to 0.1, 0.2 and 0.3 for 10%, 20% and 30% relative sampling rates, respectively; the lower bound on x_2 was set to 0.07 to ensure there is sufficient number of samples for accurate frame reconstruction.

To help facilitate training an object detection network for the bounding box detection and object classification, the authors of [15] released a dataset with 300 hours of radar data with object annotations. The authors of [15] proposed using Faster R-CNN network for predicting bounding boxes of vehicles directly on the radar frame. In Faster R-CNN, the features from the images are first processed through a region proposal network; the generated proposals are then used for the bounding box and class prediction. As an alternative, we also considered a YOLO-like model [20] trained on radar data without a region proposal step, enabling direct prediction from images split into $S \times S$ cells; such a model has the advantage of fewer parameters and thus require less computational resources while leading to faster inference. Table 3 reports the latency and model information, providing guidance on the trade-off between performance and latency. Furthermore, although various object detection networks have been studied in the context of image classification, our work is the first to conduct such a study in the context of radar object detection using the YOLO object detection model. Therefore, we trained two variants of YOLO, the medium and the large model to show the performance of radar object detection models trained on the RADIATE dataset.

TABLE 3. The AP/AP50, the number of parameters and the inference time (NVIDIA 2070 super GPU) for the three models on the RADIATE dataset.

Model	AP50	AP	Params	Inference [ms]
Faster R-CNN [15]	57.2	22.6	41M	63ms/img
YOLOv5m [20]	61.9	26.4	21M	41ms/img
YOLOv5l [20]	**63.8**	**27.0**	46M	70ms/img

For our experiments, we trained a Faster R-CNN using the parameters recommended by [15] and two YOLO models of different complexity, aiming to identify the best network in terms of performance and the total number of parameters. The Faster R-CNN network has the ResNet-50 backbone and is trained for 90,000 iterations with 128 images per batch and 0.00025 learning rate. The network was trained for a single class classification of vehicles, i.e., a collection of all released annotations except for pedestrians and groups of pedestrians. The YOLO networks [20] were also trained for vehicle prediction using radar frames of size 1280x1280. The YOLO5l network was trained for 10 epochs with batch size 2; YOLO5l has 46M parameters, performing inference on NVIDIA 2070 Super GPU took 70ms/img. The YOLO5m network was also trained for 10 epochs with batch size 2; that network had 20M parameters, the inference speed was 41ms/img. The performance of the networks on the test set is shown in Table 3.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043

Among the three models trained on the RADIATE dataset, YOLO5l achieved the best AP50 and AP but has the highest number of parameters and longest inference time. At half the number of parameters, YOLO5m required a little more than half the inference time of YOLO5l to achieve 61.9 AP50 on the entire dataset. To our knowledge, these are the first application of YOLO networks on the RADIATE data.

The performance comparison between our proposed CompRPD and baselines is reported in Table 4. Here Radar-Detect (RD) denotes the method in [22] utilizing an inverted T-shaped sampling pattern. All the methods were tested using both Faster R-CNN and YOLO networks. CompRPD (YOLOv5l) and CompRPD (Faster R-CNN) denotes the CompRPD algorithm that deployed the YOLOv5l and Faster R-CNN networks for object detection, respectively. The Faster R-CNN achieved AP50 and AP of 58.4 and 23.8, respectively, while YOLOv5l achieved AP50 and AP of 62.4 and 25.3, respectively on the original tested radar data (40 frames from each of the 5 scenes). At 20% and 30% sampling rates, the algorithms that deployed YOLO for object detection performed better than their counterparts deploying Faster R-CNN; however, at 10% sampling rate, RD and CompRPD deploying the Faster R-CNN network outperformed RD and CompRPD deploying YOLOv5. Overall, the accuracy at a 10% sampling rate deteriorated significantly from the full rate performance and thus such ultra-low sampling rates are not recommended in practice. At 20% relative sampling rate, CompRPD (YOLOv5) performed the best, achieving 55.6% AP50 and 24.1% AP. Across all sampling rates and networks, the RD and CompRPD methods surpassed the standard CS and CFAR algorithms in end-to-end object detection tasks.

The results of applying different schemes to the radar frame reconstruction are shown in Figure 3. The first row shows the reconstruction of a frame from the city scene. Among the considered schemes, CompRPD had the most predictions aligning with the ground truth. Overall, schemes deploying YOLO networks outperformed those deploying Faster R-CNN. The second row shows the reconstruction of a frame from the fog scene. Yet again, the CompRDP (YOLO) algorithm achieves better-bounding box alignment with the ground truth and, in certain cases, fewer false positive predictions.

Figure 4 shows the results of running the Kalman filter's prediction based on the previous frame information and the detection network's output on the sub-sampled and reconstructed frame using the RD (YOLOv5) algorithm. This is illustrated across 5 frames of the city and fog scenes. The Kalman filter may in certain frames assign to an object a smaller bounding box than appropriate (e.g., frame 3 of the city scene) but in the subsequent frames it recovers from this error; one can make a similar observation in the case of the object in the front (top) of the ego-vehicle (frames 4 and 5). As for the fog scene, the Kalman filter consistently accurately predicts the bounding box of the objects. In general, the use of the Kalman filter's predictions instead of the network's bounding box from the previous sub-sampled frame helps localize areas where a higher sampling rate needs to be allocated.

We further tested the improvement that the prediction of future occlusions by the objects detected in the current frame provides over dropping the occluded blocks; CompRPD (Yolov5) achieves AP50 and AP of 55.6 and 24.1, respectively, while simply dropping the occluded blocks results in AP50 and AP of 43.8 and 18.9, respectively (20% sampling rate).

TABLE 4. The AP/AP50 object detection metrics are presented for the A-1 [25], CFAR [29], Radar-Detect (RD) [22] and CompRPD algorithms using both the Faster R-CNN and YOLO networks. CompRPD outperforms baseline techniques across all the sampling rates; the optimal trade-off between performance and the sampling rate is noted at 20% sampling rate.

Sampling rate	10%	20%	30%
A-1 (Faster R-CNN) [25]	10.0/3.6	37.7/15.0	49.2/20.3
CA-CFAR (Faster R-CNN) [29]	11.9/4.5	37.7/14.3	48.3/19.6
GO-CFAR (Faster R-CNN)	10.4/3.9	36.5/14.5	49.0/20.2
RD (Faster R-CNN) [22]	35.5/14.9	51.9/21.4	55.0/22.9
CompRPD (Faster R-CNN)	34.3/14.7	**53.4/21.9**	55.0/22.8
A-1 (YOLOv5) [25]	7.1/2.4	36.3/14.2	54.3/22.0
CA-CFAR (YOLOv5) [29]	8.0/3.0	37.1/13.9	51.5/21.0
GO-CFAR (YOLOv5)	6.7/2.5	36.1/13.8	52.3/21.0
RD (YOLOv5) [22]	27.9/12.6	54.6/23.8	61.4/26.3
CompRPD (YOLOv5)	27.9/12.7	**55.6/24.1**	61.4/26.3

V. Conclusion

We presented a radar sub-sampling pipeline that relies on prior images and/or radar data (CompRADIMG) to accomplish effective reconstruction using as few as 10% of the samples. In the case of the sub-sampling algorithm that uses only the prior radar data (CompRPD), we showed effective reconstruction using as few as 20% of the sub-sampled radar data. The CompRADIMG algorithm was tested on the open-sourced Oxford Range-Azimuth radar data taken during good weather conditions, while the CompRPD algorithm was tested on the radiate radar data acquired in fog, snow, motorway, night and city conditions. We further reported results on using hardware-efficient BPD measurement matrices for compressed sensing, showing that on the considered automotive data they outperform BPBD and Gaussian measurement matrices. Finally, we presented a YOLO-based radar-object detection network and showed a significant performance improvement over its Faster R-CNN based counterpart.

REFERENCES

[1] Felix Nobis, Maximilian Geisslinger, Markus Weber, Johannes Betz, and Markus Lienkamp, "A deep learning-based radar and camera sensor fusion architecture for object detection," in *2019 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, 2019, pp. 1–7.

[2] 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,"

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043.

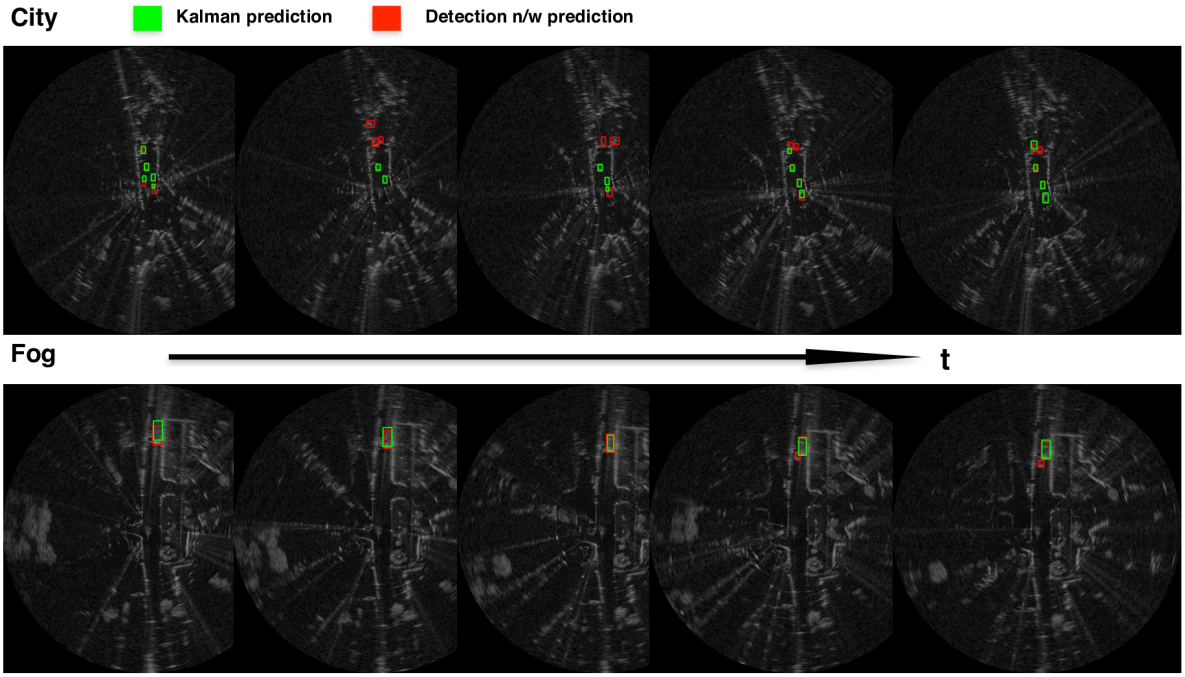


FIGURE 4. Consecutive frames are shown from the city and the fog dataset. The green box illustrates Kalman's predictions based on the previous frame and the red box illustrated the current frame's bounding boxes predictions. After every Kalman update across frames, indicated by the frame numbers the Kalman predictions are closer to the location of the object shown by the bounding boxes.

IEEE Journal of Selected Topics in Signal Processing, vol. 15, no. 4, pp. 954–967, 2021.

[3] Ramin Nabati and Hairong Qi, "Centerfusion: Center-based radar and camera fusion for 3d object detection," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, January 2021, pp. 1527–1536.

[4] M. Meyer and G. Kuschk, "Deep learning based 3d object detection for automotive radar and camera," in *2019 16th European Radar Conference (EuRAD)*, 2019, pp. 133–136.

[5] Shuo Chang, Yifan Zhang, Fan Zhang, Xiaotong Zhao, Sai Huang, Zhiyong Feng, and Zhiqing Wei, "Spatial attention fusion for obstacle detection using mmwave radar and vision sensor," *Sensors*, vol. 20, no. 4, 2020.

[6] Ramin Nabati and Hairong Qi, "Rrpn: Radar region proposal network for object detection in autonomous vehicles," in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 3093–3097.

[7] Saptarshi Mukherjee, Andrew M. Wallace, and Sen Wang, "Predicting vehicle behavior using automotive radar and recurrent neural networks," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 2, pp. 254–268, 2021.

[8] Rui Chen, Jiameng Ning, Yu Lei, Yilong Hui, and Nan Cheng, "Mixed traffic flow state detection: A connected vehicles-assisted roadside radar and video data fusion scheme," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 4, pp. 360–371, 2023.

[9] Ana Stroescu, Liam Daniel, and Marina Gashinova, "Combined object detection and tracking on high resolution radar imagery for autonomous driving using deep neural networks and particle filters," in *2020 IEEE Radar Conference (RadarConf20)*, 2020, pp. 1–6.

[10] Chih-Chung Hsu, Chieh Lee, Lin Chen, Min-Kai Hung, Yu-Lun Lin, and Xian-Yu Wang, "Efficient-rod: Efficient radar object detection based on densely connected residual network," in *Proceedings of the 2021 International Conference on Multimedia Retrieval*, New York, NY, USA, 2021, ICMR '21, p. 526–532, Association for Computing Machinery.

[11] Pengliang Sun, Xuetong Niu, Pengfei Sun, and Kele Xu, "Squeeze-and-excitation network-based radar object detection with weighted location fusion," New York, NY, USA, 2021, ICMR '21, p. 545–552, Association for Computing Machinery.

[12] Shoaib Azam, Farzeen Munir, and Moongu Jeon, "Channel boosting feature ensemble for radar-based object detection," in *2021 IEEE Intelligent Vehicles Symposium (IV)*, 2021, pp. 762–769.

[13] Bo Ju, Wei Yang, Jinrang Jia, Xiaoqing Ye, Qu Chen, Xiao Tan, Hao Sun, Yifeng Shi, and Errui Ding, "Danet: Dimension apart network for radar object detection," in *Proceedings of the 2021 International Conference on Multimedia Retrieval*, New York, NY, USA, 2021, ICMR '21, p. 533–539, Association for Computing Machinery.

[14] Ana Stroescu, Liam Daniel, Dominic Phippen, Mikhail Cherniakov, and Marina Gashinova, "Object detection on radar imagery for autonomous driving using deep neural networks," in *2020 17th European Radar Conference (EuRAD)*, 2021, pp. 120–123.

[15] 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)*, 2021, pp. 1–7.

[16] R.H. Walden, "Analog-to-digital converter survey and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539–550, 1999.

[17] Emmanuel J Candès and Michael B Wakin, "An introduction to compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[18] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 03 1960.

[19] Yasser H. Khalil and Hussein T. Mouftah, "Licanet: Further enhancement of joint perception and motion prediction based on multi-modal fusion," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 222–235, 2022.

[20] Glenn Jocher, Ayush Chaurasia, Alex Stoken, Jirka Borovec, NanoCode012, Yonghye Kwon, TaoXie, Kalen Michael, Jiacong Fang, imyhxy, Lorna, Colin Wong, Zeng Yifu, Abhiram V, Diego Montes, Zhiqiang Wang, Cristi Fati, Jebastin Nadar, Laughing, UnglvKitDe, tkianai, yxNONG, Piotr Skalski, Adam Hogan, Max Strobel, Mrinal Jain, Lorenzo Mammana, and xylieong, "ultralytics/yolov5: v6.2 - YOLOv5 Classification Models, Apple M1, Reproducibility, ClearML and Deci.ai integrations," Aug. 2022.

[21] Dan Barnes et al., "The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset," *arXiv*, 2019.

[22] Madhumitha Sakthi, Ahmed Tewfik, Marius Arvinte, and Haris Vikalo, "End-to-end system for object detection from sub-sampled radar data," 2022.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

This article has been accepted for publication in IEEE Open Journal of Intelligent Transportation Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/OJITS.2023.3332043

[23] Madhumitha Sakthi and Ahmed Tewfik, "Automotive radar data acquisition using object detection," in *2021 29th European Signal Processing Conference (EUSIPCO)*, 2021, pp. 1770–1774.
[24] Seongwook Lee, Yunho Jung, Myeongjin Lee, and Wookyung Lee, "Compressive sensing-based sar image reconstruction from sparse radar sensor data acquisition in automotive fmcw radar system," *Sensors*, vol. 21, no. 21, 2021.
[25] Z. Slavik et al., "Compressive sensing-based noise radar for automotive applications," in *2016 12th IEEE International Symposium on Electronics and Telecommunications (ISETC)*, 2016, pp. 17–20.
[26] Fabian Roos, Philipp Hügler, Christina Knill, Nils Appenrodt, Jürgen Dickmann, and Christian Waldschmidt, "Data rate reduction for chirp-sequence based automotive radars using compressed sensing," in *2018 11th German Microwave Conference (GeMiC)*, 2018, pp. 347–350.
[27] A. Correas-Serrano and M. A. González-Huici, "Experimental evaluation of compressive sensing for doa estimation in automotive radar," in *2018 19th International Radar Symposium (IRS)*, 2018, pp. 1–10.
[28] S. K. Sahoo and A. Makur, "Signal recovery from random measurements via extended orthogonal matching pursuit," *IEEE Transactions on Signal Processing*, vol. 63, no. 10, pp. 2572–2581, 2015.
[29] A. M. Assem, R. M. Dansereau, and F. M. Ahmed, "Adaptive sub-nyquist sampling based on haar wavelet and compressive sensing in pulsed radar," in *2016 4th International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, 2016, pp. 173–177.
[30] I. Kyriakides, "Adaptive compressive sensing and processing for radar tracking," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 3888–3891.
[31] J. Zhang, D. Zhu, and G. Zhang, "Adaptive compressed sensing radar oriented toward cognitive detection in dynamic sparse target scene," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, 2012.
[32] Zhihui Cao, Junjie Li, Chunyi Song, Zhiwei Xu, and Xiaoping Wang, "Compressed sensing-based multitarget cfar detection algorithm for fmcw radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 11, pp. 9160–9172, 2021.
[33] Xuan Truong Nguyen, Khac-Thai Nguyen, Hyuk-Jae Lee, and Hyun Kim, "Roi-based lidar sampling algorithm in on-road environment for autonomous driving," *IEEE Access*, vol. 7, pp. 90243–90253, 2019.
[34] Feng Xi, Nir Shlezinger, and Yonina C. Eldar, "Bilimo: Bit-limited mimo radar via task-based quantization," *IEEE Transactions on Signal Processing*, vol. 69, pp. 6267–6282, 2021.
[35] Aria Ameri, Arindam Bose, Jian Li, and Mojtaba Soltanalian, "One-bit radar processing with time-varying sampling thresholds," *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5297–5308, 2019.
[36] Feng Xi, Yijian Xiang, Zhen Zhang, Shengyao Chen, and Arye Nehorai, "Joint angle and doppler frequency estimation for mimo radar with one-bit sampling: A maximum likelihood-based method," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 6, pp. 4734–4748, 2020.
[37] Z. He, "The simplest measurement matrix for compressed sensing of natural images," in *2010 IEEE International Conference on Image Processing*, 2010, pp. 4301–4304.
[38] A. Akbari, "Robust image reconstruction for block-based compressed sensing using a binary measurement matrix," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018.
[39] Xianxiang et al. Chen, "Compressed sensing based method for electrocardiogram monitoring on wireless body sensor using binary matrix," *International Journal of Wireless and Mobile Computing*, vol. 8, no. 2, pp. 114–121, 2015.
[40] A. Ravelomanantsoa, "Compressed sensing: A simple deterministic measurement matrix and a fast recovery algorithm," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 12, pp. 3405–3413, 2015.
[41] L. Miccinesi et al., "Compressive sensing for no-contact 3d ground penetrating radar," in *2018 41st International Conference on Telecommunications and Signal Processing (TSP)*, 2018, pp. 1–5.
[42] F. Parvaresh, H. Vikalo, S. Misra, and B. Hassibi, "Recovering sparse signals using sparse measurement matrices in compressed dna microarrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 3, pp. 275–285, 2008.
[43] Shuo Chang et al., "Spatial attention fusion for obstacle detection using mmwave radar and vision sensor," *Sensors (Basel, Switzerland)*, vol. 20, no. 4, February 2020.
[44] R. Nabati and H. Qi, "Rrpn: Radar region proposal network for object detection in autonomous vehicles," in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 3093–3097.
[45] Holger Caesar et al., "nuscenes: A multimodal dataset for autonomous driving," *arXiv*, 2020.
[46] Zhi Tian, Chunhua Shen, Hao Chen, and Tong He, "Fcos: Fully convolutional one-stage object detection," 2019.
[47] Ali Akbari and Maria Trocan, "Robust image reconstruction for block-based compressed sensing using a binary measurement matrix," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 1832–1836.
[48] David L Donoho and Michael Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.
[49] David L Donoho and Michael Elad, "On the stability of the basis pursuit in the presence of noise," *Signal Processing*, vol. 86, no. 3, pp. 511–532, 2006.
[50] Emmanuel Candes and Justin Romberg, "l1-magic: Recovery of sparse signals via convex programming," *URL: www. acm. caltech. edu/l1magic/downloads/l1magic. pdf*, vol. 4, no. 14, pp. 16, 2005.
[51] Irfan Mehmood, Ran Li, Xiaomeng Duan, Xiaoli Guo, Wei He, and Yongfeng Lv, "Adaptive compressive sensing of images using spatial entropy," *Computational Intelligence and Neuroscience)*, 2017.
[52] Mark Richards, "Fundamentals of radar signal processing," 2005.
[53] V. Hansen, "Constant false-alarm-rate processing in search radars," *Proc. of IEEE Radar Conference*, pp. 323–325, 1973.



Madhumitha Sakthi (Student Member, IEEE) received her B.Tech in Electronics and Communication engineering from SASTRA University, India in 2017 and MSE and Ph.D. in Electrical and Computer Engineering from The University of Texas at Austin in 2019 and 2022 respectively. Since 2022, she is a Video Research Engineer with Qualcomm, USA. Her research interests include sensor data fusion for autonomous driving applications and deep learning for signal processing.



Marius Arvinte received his B.S. and M.S. degrees in Electronics, Telecommunications and Information Technology from the Politehnica University of Bucharest, Romania in 2014 and 2016, respectively, and the Ph.D. degree in Electrical and Computer Engineering from the University of Texas at Austin, USA in 2022. Since 2022, he is a Research Scientist with Intel Labs, USA. His research interests include applied deep learning for signal processing and inverse problems, with a focus on generative models.



Haris Vikalo (Senior Member, IEEE) received the B.S. degree from the University of Zagreb, Croatia, in 1995, the M.S. degree from Lehigh University in 1997, and the Ph.D. degree from Stanford University in 2003, all in electrical engineering. He held a short-term appointment at Bell Laboratories, Murray Hill, NJ, in the summer of 1999. From January 2003 to July 2003 he was a Postdoctoral Researcher, and from July 2003 to August 2007 he was an Associate Scientist at the California Institute of Technology. Since September 2007, he has been faculty at the Department of Electrical and Computer Engineering, the University of Texas at Austin. He is a recipient of the 2009 National Science Foundation Career Award. His research interests include signal processing, machine learning, communications and bioinformatics.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/