Improving Millimeter Wave Radar Perception with Deep Learning

Junfeng Guan Sohrab Madani
ECE 544 Project Reportl
jguan8@illinois.edu smadani2@illinois.edu

1. Introduction and Project Descriptions

Since the past few years, AI-Powered autonomy revolution in the automotive industry has attracted great attention worldwide. It is believed that in the not-too-distant future, fullyautonomous vehicles will be the norm rather than the exception, redefining mobility in our daily lives. With deep learning widely applied on sensor data, self-driving cars are able to localize and map objects, understand the environment, and make correct decisions. As the most fundamental task, previous works have demonstrated accurate object detection and classification, but they are limited to data obtained from LiDARs and cameras. These optical sensors have high imaging resolution, but they naturally fail in low visibility conditions such as fog, rain, and snow, because light beams are narrower than water droplets and snowflakes.[?] This fundamental limitation of optical sensors is one of the major roadblocks to achieving the 5th SAE level of full automation. [1] On the contrary, Radar wave propagates through small particles and can provide an alternate imaging solution in such inclement weather. Besides, radar can also directly measure the velocity of vehicles based on the doppler shift of the reflected signal instead of going through cluster tracking among frames like LiDAR. With the additional realtime speed information, AI can potentially have a better perception than human drivers.

Although the low resolution of traditional automotive radar overshadows its advantages, the advent of Millimeter-wave (mmWave) technology makes it possible to have a reliable 3D imaging system in inclement weather with relatively higher resolution. Along with good propagation characteristics, mmWave also provides much wider bandwidth and enables miniature-sized large-aperture antenna arrays, which improve distance and angle measurement respectively. Previous work has demonstrates sub-centimeter level imaging resolution for short range objects [4], but the fundamental difference in wavelength between radar (5 mm) and LiDAR (905 nm) cannot be easily overcome. Moreover, radar images are not as readable to non-experts. Therefore, in order to convince the automotive industry and general public that mmWave is a reliable imaging solution for autonomous driving, we should try to generate radar images targeting the resolution comparable to LiDAR, and make them more perceptually intuitive to people.

In this project, we propose to develop high-resolution and reliable imaging techniques for self-driving cars with mmWave radar. Specifically training neural networks to enhance the low-resolution and unreadable radar images to be similar to the LiDAR point cloud, which has been extensively and successfully used for self-driving perception, and eventually enable various crucial vision applications for autonomous vehicles like lane detection, image mapping, localization, and object identification with mmWave radar data only.

2. Challenges

2.1. Radar Imaging Primer

Radar generates images of objects by localizing the cluster of point reflectors that forms the object. 3D space imaging requires mapping point reflectors to voxels in a spherical coordinates with its distance, azimuth angle, and elevation angle. Distance is measured through to round-trip Time-of-Flight (ToF) of reflected radar signal, while azimuth and elevation angles can be either obtained by the beam steering angle of phased array antennas or be estimated with Direction-of-Arrival (DoA) estimation algorithms such as beam-forming or Multiple Signal Classification (MUSIC). Notice that as distance becomes further, the voxel size corresponds to the same angle increases, so that long-range objects contain much fewer voxels and appear to be more blurry. Besides, unlike the extremely narrow width of light beams, the cone-shape radar beam with sidelobes cause interference and leakage among nearby voxels and even the environment, which smears generated images. Last but not least, with a few centimeter resolution of distance measurement, a continuum of a large number of point reflector sums up and makes it an under-determined problem to localize them. Besides, radar reflection tends to be more specular than the mostly scattering reflection at optical frequencies. In other words, reflection off smooth objects might mainly towards an angle away from the radar receiver and disappears in the image. Hence, ”edge detection” in radar images is very different from that of pictures. Firstly, it needs to predict high frequency information with only low frequency data. Secondly, it needs to learn to fill up missing parts of the object.

2.2. Related Work

Previous works have attempted to adapt the optical camera-oriented Convolutional Neural Network (CNN) to its microwave counterpart to classify single-object images from high-resolution 2D synthetic aperture radar (SAR) images. [2] [6] [3] There has also been successful application of CNNs on recreating short range human body skeletons in 2D images and 3D spaces from radar images by tracking 14 key points. [7] [8] However, the scope of these application of neural networks on radar images are restricted to single object classification and short-range human body parts detection and classification. As our resolution degrade.

2.3. Model Selection

Therefore well studies CNN structure -¿ GAN.

2.4. Dataset Availability

Once the network is setup, it needs training data -i.e., it needs many labeled exp Dataset Variation between systems Experiment Processing

2.5. 3D Complexity

3.3D 3D CNN 3D GAN size complexity

2.6. Evaluation Metrics

3. Method

Describe the overall method on how you solve the proposed problem, and a bit of original derivation that has some relevance to what you're trying to accomplish

Problem statement: low blured images no boundary can be seen, specularity causes missing parts, no availbale dataset, 4D CNN NN complex.

Overview 3D, in the method, we say that we start with 2D version, and based on that build 3D. We are trying to use cGAN to generate higher resolution images from low resolution radar images, which should have a sharp and accurate boundary of the object. Also, the missing parts due to specuality of reflection need to be feel up. S

cGAN: Why cGAN:

GAN is tranfering images from 1 domain to another, some application

Generative Adverserial Networks (or GANs) have been widely used to generate images [cite some stuff here] and fill in missing parts of data. In our case, we are looking to generate images with accurate boundaries using low resolution images with missing parts as input. Conditional GANs have already proven successful in similar settings, such as in [5], where the authors have used thermal images under low light conditions where there some parts of the image are missing to retrieve the human face boundaries and estimate its orientation. Another motivation behind using GANs is that generic loss functions such as the L_2 and L_1 (i.e. A loss function that tries to minimize the Euclidean or L_1 distance between the input and the output) which are the de facto standard loss function for restoring images render blurry images, which are not suitable for our application. On the other hand, we try to motivate the loss in GAN to learn to focus on the boundaries, by designing the ground to contain information mostly about the boundary of objects, as we shall see in the dataset section CGANs take as input a Transition: cGAN has input, groundtruth, G, D

Input: Low resolution radar images, because there is no available public dataset of mmWave radar images, and collecting a big enough dataset by ourselves is not possible. We synthesized radar images. This heat-map is fed into the network. Groundtruth: 1. Size of 3D which makes the training phase very slow even when using GPUs. 2. 2D: 3. 3D

The input to our problem is pre-processed radar data. First, using the raw data from the antenna array, a coarse heat-map of objects are generated. After some further processing.

Why not raw data: (Goes to detail of Dataset jayden) Radar imaging processing algorithms is a well establish feild for 70 years. and there is fruitless to try and learn them using machine learning. So instead of The reason why we did not choose the raw data as input is twofold.

3.1. Conditional GAN

3.2. Dataset Generation

Once we have designed our Unlike the digital camera, imaging radar less common, no dataset available . Since there is no

Challenge of unavailable radar dataset.

Experiment to collect radar images size time umbiguity

Simulation with EM

3D

2D Mask R-CNN input: radar image groundtruth: mask pros: large dataset with car truck human cons: No 3D info, no specularity 3D CAD input: contour groundtruth: pros: small dataset, single element cons: 3D shape info, specularity

Evaluate the simulation with EM simulation Feko and experimental results.

4. Experimental Results

Describe the setup of the experiments you ran, e.g., what evaluation metrics, datasets are used. Present the results, preferably in the form of tables and/or figures

4.1. Dataset

4.2. Results

5. Discussion and Conclusion

Analyze the results, summarize the findings and point out possible future directions

References

[1] Automated vehicles for safety. 1

[2] S. Chen and H. Wang. Sar target recognition based on deep learning. *International Conference on Data Science and Advanced Analytics, DSAA*, 2014. 2

[3] M. Gong, J. Zhao, J. Liu, Q. Miao, and L. Jiao. Change detection in synthetic aperture radar images based on deep neural networks. *IEEE Transactions on Neural Networks and Learning Systems , DSAA*, 2016. 2

[4] B. Mamandipoor, G. Malysa, A. Arbabian, U. Madhow, and K. Noujeim. 60 ghz synthetic aperture radar for short-range imaging: Theory and experiments. *Asilomar Conference on Signals, Systems and Computers*, 2014. 1

[5] A. Nambi, S. Bannur, I. Mehta, H. Kalra, A. Virmani, V. Padmandabhan, R. Bhandari, and B. Raman. Demo: Hams: Driver and driving monitoring using a smartphone. *ACM MobiCom*, 2018. 2

[6] C. Schwegmann, W. Kleynhans, B. Salmon, L. Mdakane, and R. Meyer. Very deep learning for ship discrimination in synthetic aperture radar imagery. *IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, 2016. 2

[7] M. Zhao, T. Li, M. Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi. Trhough-wall human pose estimation using radio signals. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2018. 2

[8] M. Zhao, Y. Tian, H. Zhao, M. Alsheikh, T. Li, R. Hristov, Z. Kabelac, D. Katabi, and A. Torralba. Rf-based 3d skeletons. *SIGCOMM*, 2018. 2