

An Adversarial Super-Resolution Remedy for Radar Design Trade-offs

Karim Armanious^{1*}, Sherif Abdulatif^{1*}, Fady Aziz², Urs Schneider² and Bin Yang¹

¹Institute of Signal Processing and System Theory, University of Stuttgart, Germany

²Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Stuttgart, Germany

Email: {karim.armanious, sherif.abdulatif}@iss.uni-stuttgart.de

*These authors contributed to this work equally.

Abstract—Radar is of vital importance in many fields, such as autonomous driving, safety and surveillance applications. However, it suffers from stringent constraints on its design parametrization leading to multiple trade-offs. For example, the bandwidth in FMCW radars is inversely proportional with both the maximum unambiguous range and range resolution. In this work, we introduce a new method for circumventing radar design trade-offs. We propose the use of recent advances in computer vision, more specifically generative adversarial networks (GANs), to enhance low-resolution radar acquisitions into higher resolution counterparts while maintaining the advantages of the low-resolution parametrization. The capability of the proposed method was evaluated on the velocity resolution and range-azimuth trade-offs in micro-Doppler signatures and FMCW uniform linear array (ULA) radars, respectively.

Index Terms—Radar, Super-resolution, Micro-Doppler, MIMO, Range-azimuth, Convolutional neural network, CNN, Generative adversarial networks, GAN, Remote sensing

I. INTRODUCTION

Radar is one of the most powerful tools for environment sensing. This is due to its superior capabilities in harsh environments and under low lighting conditions, where cameras and other vision-based sensors usually fail to operate. However, the design of a radar system for a particular application is often subject to some practical constraints and trade-offs between different system parameters [1]. For instance, the bandwidth is inversely proportional to both the range resolution and the maximum unambiguous detectable range. Thus, having a higher range resolution will lead to a deterioration in the maximum unambiguous detectable range. The same applies for the velocity. Faster chirps will enhance the velocity resolution at the expense of a lower maximum detectable velocity [2]. For MIMO radars, increasing the number of antennas will enhance the angular resolution but the longer acquisition time impairs real-time applications [3].

Accordingly, radar systems experience a fundamental limitation in many applications and they can not be optimum in all properties simultaneously. As a result, radar is often used in fusion with other sensors and not as a standalone sensor. For example, autonomous driving is the most actively researched area where radar is utilized in addition to other electro-optical sensors such as camera and LIDAR [4]. These sensors can either operate independently or co-dependently to construct 3D models of the surroundings. Recently, radar is also employed for indoor monitoring applications due to its insensitivity against smoke and lighting conditions. For instance, a fusion between radar and thermal infrared camera (TIR) is proposed in [5] to detect real hostages from mirrored TIR reflections in a fire-fighting scenario.

Additionally, radar imaging applications are also susceptible to radar design trade-offs. In these applications, a scene is constructed as a 2D or 3D image using either electronic beam steering [3] or mechanical scanning, such as synthetic aperture radar (SAR) [6] or inverse synthetic aperture radar (ISAR) [7]. Active research in radar imaging leads to the evolution of different areas and applications such as through wall imaging [8], land and weather monitoring [9]. However, these techniques suffer from trade-offs between their spatial parameters (range-azimuth-elevation resolutions) due to design constraints on bandwidth, acquisition time, number of antennas and antenna separation.

Several methods have been proposed to enhance the acquired radar measurements. This family of methods is called radar super-resolution. They operate in general on the collected IQ time-series. In [10], the authors used finer sampling to enhance resolution of both range and azimuth in NEXRAD weather data. They also utilized SZ-Phase coding on the time series data to enhance the Doppler resolution on the same dataset [11]. Super-resolution of SAR and ISAR imaging is also extensively studied by treating the problem as sparse signal reconstruction based on an orthogonal basis. Then convex optimization or Bayesian deconvolution are used to solve the reconstruction problem [12], [13]. A very similar approach is applied in [14] to enhance the resolution of direction of arrival (DOA) estimation using MUSIC algorithm.

In this work, we propose a new technique for radar super-resolution to bypass the aforementioned practical constraints and trade-offs regarding radar parametrization. The proposed technique operates on the post-processed 2D time-frequency or range-azimuth representations/images. Unlike the previous approaches operating on time-series IQ data, this more accessible image-based approach allows the utilization of state-of-the-art techniques from the deep learning community to achieve radar super-resolution.

In recent years, the significant increase in data size and computational resources leads to a considerable boost in deep learning. Especially the development of deep convolutional neural networks (DCNNs) radically changed the research approach in the field of image processing. Generative adversarial networks (GANs), introduced by Ian Goodfellow in [15], are a recent branch and are considered as the state-of-the-art in image generation tasks. Moreover, a variant of GANs, conditional generative adversarial networks (cGANs) [16], was recently introduced for image-to-image translation such as natural image super-resolution [17], medical image correction and inpainting [18], [19].

In this paper, we utilize a modified cGAN framework for the task of radar super-resolution. The proposed framework incorporates additional non-adversarial losses to enhance the details of the super-resolved radar images. To validate the effect of the proposed framework, experiments were conducted on the super-resolution of micro-Doppler (μ -D) signatures of walking human targets. Additionally, some preliminary experiments on the super-resolution of range-azimuth spectrum in FMCW uniform linear array (ULA) radar were conducted.

II. EXPERIMENTAL SETUP

A. Micro-Doppler Signatures

The μ -D signature of a moving target is identified by the time-frequency (TF) analysis of the backscattered IQ signal. The frequency shifts in the down-converted backscattered signal represent the velocity information of the specific target. It can be used to visualize the velocity variation of a certain target of interest over time. Accordingly, a moving rigid body, such as pendulum, will have only one velocity component which can be either varying or constant. On the other hand, complex bodies, such as humans, will have different velocity components representing the motion of each limb during a gait cycle [20].

In the conducted experiments, the proposed adversarial framework for radar super-resolution will be applied on the μ -D signatures of walking humans. The main motivation is that the μ -D signatures contain rich identifying information. Thus, they can be used to fingerprint a walking person based on unique walking style and body mass index (BMI) [21]. As such, the proposed radar super-resolution framework aims at reconstructing a high-resolution μ -D signature from a low-resolution one while preserving the identity information of each person.

The TF representation is derived from the short-time Fourier transform (STFT) of the complex time-domain baseband signal. In our method, a Gaussian window with a fixed window size of 512 and a window overlap of 75% is used to obtain all μ -D signatures. A CW radar operating at a carrier frequency of $f_c = 25$ GHz is used for data collection. Based on the relations presented in [20], the maximum unambiguous velocity (v_{max}) and the corresponding velocity resolution (v_{res}) of the TF representation are

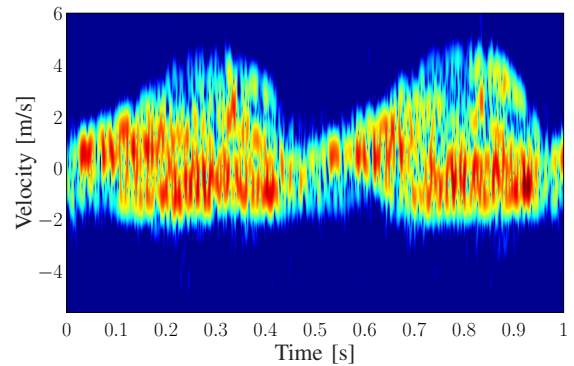
$$v_{max} = \frac{cf_p}{4f_c} \quad \text{and} \quad v_{res} = \frac{2v_{max}}{w} \quad (1)$$

where c is the speed of light, f_p is the pulse repetition frequency and w is the STFT window size.

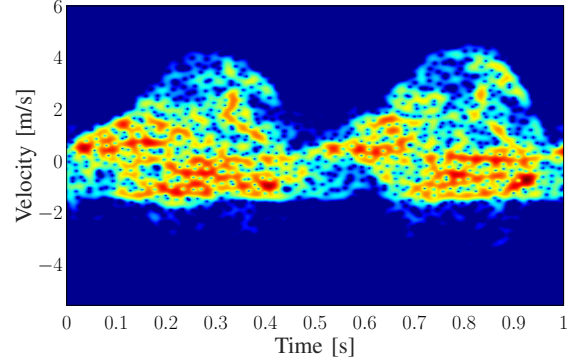
A clear trade-off exists between the maximum unambiguous velocity range and the resultant velocity resolution. An enhancement of v_{max} will lead to a corresponding deterioration of the resolution and vice versa. In this work, we aim to overcome this radar design trade-off by translating a low-resolution μ -D signature to a corresponding high-resolution signature while keeping the same window size and hence the STFT complexity.

B. Dataset Preparation and Collection

In our experiments, we obtained low-resolution and high-resolution μ -D signatures by tuning the f_p parameter based on Eq. 1. For the high-resolution case, f_p is set to 2 kHz to



(a) Low resolution μ -D.



(b) High resolution μ -D.

Fig. 1: An example of a paired high and low-resolution μ -D signatures of one full gait cycle for the same subject.

obtain a maximum velocity of 6 m/s and a velocity resolution of 2 cm/s. In the low-resolution experiment, f_p is increased to 8 kHz and accordingly the maximum velocity is quadrupled to 24 m/s with a lower velocity resolution of 8 cm/s. The f_p can be controlled in offline data processing by sampling the acquired continuous-wave IQ data of a single measurement. Thus, perfectly paired low and high-resolution μ -D signatures can be acquired using a single radar acquisition, as illustrated in Fig. 1. This pixel-wise pairing of the signatures is mandatory for training the utilized adversarial framework.

For the low-resolution images, the STFT operation results in a velocity range of -24 to 24 m/s over the selected window size. Nonetheless, for the high-resolution case only a velocity range of -6 to 6 m/s is achievable. However, it was found out that a velocity of ± 6 m/s is sufficient to cover the maximum velocity of the human gait signature, as illustrated in Fig. 1. Thus, only a velocity range of -6 to 6 m/s is considered for the low-resolution images resulting in an effective window size of 128 points compared to 512 for the high-resolution case.

For dataset collection, 22 subjects of different genders, heights and weights are asked to walk on a treadmill at an intermediate velocity of 1.5 m/s. The radar is placed behind the treadmill at a fixed distance of 3 m and the direction of walking is always away from the radar. The experiment duration for each subject is 180 s and then the collected IQ time-domain signals are decimated with different decimation factors to control the f_p corresponding to both high and low-resolution signatures. Finally, the μ -D signatures of each walking subject is calculated as the STFT of the processed time signal as described above.

A full human gait cycle is divided into two swinging half gait cycles and each half gait is representing the limbs swinging velocity components for either left or right body side [22]. Since the targets are walking on a treadmill, the bulk motion of the targets is always around 0 m/s and only the swinging limbs velocity components are preserved as depicted in Fig. 1. Moreover, a clear difference between half swinging gait cycles of left and right body side is observed. Thus, the collected data is divided on half gait basis to add more variations in the dataset which will force the network to learn more detailed features. A total number of 5400 half gait signatures collected from 15 different subjects (360 cycles per subject) are used for training. The remaining unseen 7 subjects are used for testing the network performance over 2520 half gait signatures.

For training the adversarial framework, a final pre-processing step is required. The low and high-resolution signatures were resized to the same fixed size of $256 \times 256 \times 3$ unsigned 16-bit images. This was achieved by using linear interpolation to up-scale the low-resolution signatures and down-scale the high-resolution counterparts to achieve a fixed image size.

III. METHODOLOGY

The framework for radar spectrum super-resolution is based on cGANs with the inclusion of additional non-adversarial losses and a more advanced generator architecture. The proposed framework is illustrated in Fig. 2.

A. Conditional Generative Adversarial Networks

The utilized framework consists of two Deep Convolutional Neural Networks (DCNNs), together forming a conditional adversarial framework [16]. The first network, the generator G , represents a mapping function from the input low-resolution μ -D signature y to the corresponding high-resolution fake signature $\hat{x} = G(y)$. The second network, the discriminator D , receives as inputs the reconstructed high-resolution signature \hat{x} together with the target ground truth signature x . It then proceeds to classify which is the real spectrum, $D(x, y) = 1$, and which is a fake spectrum, $D(\hat{x}, y) = 0$. Both neural networks are trained against each other in competition, with the discriminator attempting to correctly classify the fake generated output and the generator attempting to improve the quality of the fake output to fool the discriminator. This training procedure is expressed as a min-max optimization task with respect to the generator and discriminator over the following adversarial loss function:

$$\min_G \max_D \mathcal{L}_{adv} = \mathbb{E}_{x,y} [\log D(x, y)] + \mathbb{E}_{\hat{x},y} [\log (1 - D(\hat{x}, y))] \quad (2)$$

B. Non-adversarial Losses

It has been discussed previously that relying solely on the adversarial loss results in translated images with visual artifacts affecting the output quality. Thus, it has been suggested to incorporate auxiliary non-adversarial losses to further regularize the generator network. Inspired by the pix2pix framework, an additional L1 loss is utilized to penalize the pixel-wise discrepancy between the translated output and the ground truth target signatures [16]. This loss is given by

$$\mathcal{L}_{L1} = \mathbb{E}_{x,y} [\|x - G(y)\|_1] \quad (3)$$

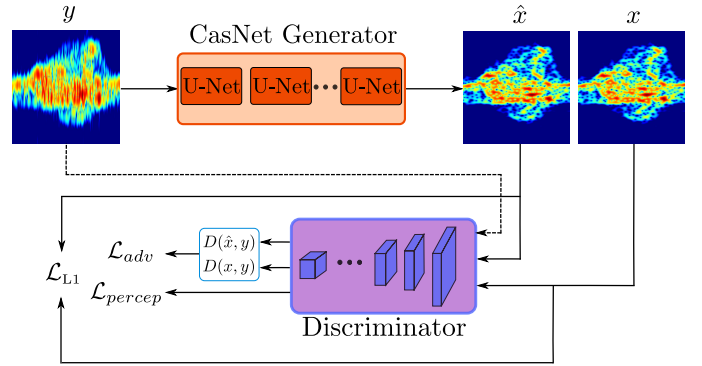


Fig. 2: An overview of the proposed adversarial architecture for μ -D signatures super-resolution.

Moreover, to avoid blurring and inconsistencies in the local structures of the results, it has been suggested by previous adversarial frameworks, such as PAN and MedGAN [23], [24], to further utilize a feature-based perceptual loss to regularize the generator network. This loss function penalize the discrepancy between intermediate feature maps extracted by the discriminator network. The perceptual loss is given by

$$\mathcal{L}_{Percep} = \sum_{i=0}^L \lambda_{pi} \|F_i(x) - F_i(\hat{x})\|_1 \quad (4)$$

where F_i is the intermediate feature map extracted from the i^{th} layer of the discriminator. L is the total number of layers and $\lambda_{pi} > 0$ is the weight contribution for each individual layer. The final loss function utilized in the adversarial framework is the sum of the previously introduced adversarial, pixel and perceptual loss functions.

C. Architectural and Implementation Details

Inspired by MedGAN, the super-resolution framework utilizes a CasNet architecture for the generator network [24]. CasNet concatenates in an end-to-end manner multiple U-Net architectures, introduced in [25], to progressively refine the translated outputs. Each U-Net architecture consists of an encoder-decoder structure with multiple skip connections to avoid loss of information due to the bottleneck layer. The utilized CasNet architecture consists of a concatenation of five U-Nets architectures. For the discriminator architecture, a patch discriminator architecture, introduced by the pix2pix framework, was utilized [16]. The patch discriminator divides the input images into 70×70 patches before classifying each of them into being real or fake images. The classification score of all image patches is then averaged out. More in-depth architectural details and consideration can be found in [24].

The framework was trained for 100 epochs using the ADAM optimizer [26] on a single NVIDIA Titan X GPU. Training time was approximately 36 hours with an inference time per image of 150 ms during validation.

IV. RESULTS

To validate the capability of the utilized adversarial framework in the super-resolution μ -D signatures, the discussion will be based on qualitative and quantitative analysis. The qualitative results are presented in Fig. 3. Through learning the joint distribution mapping between the low-resolution

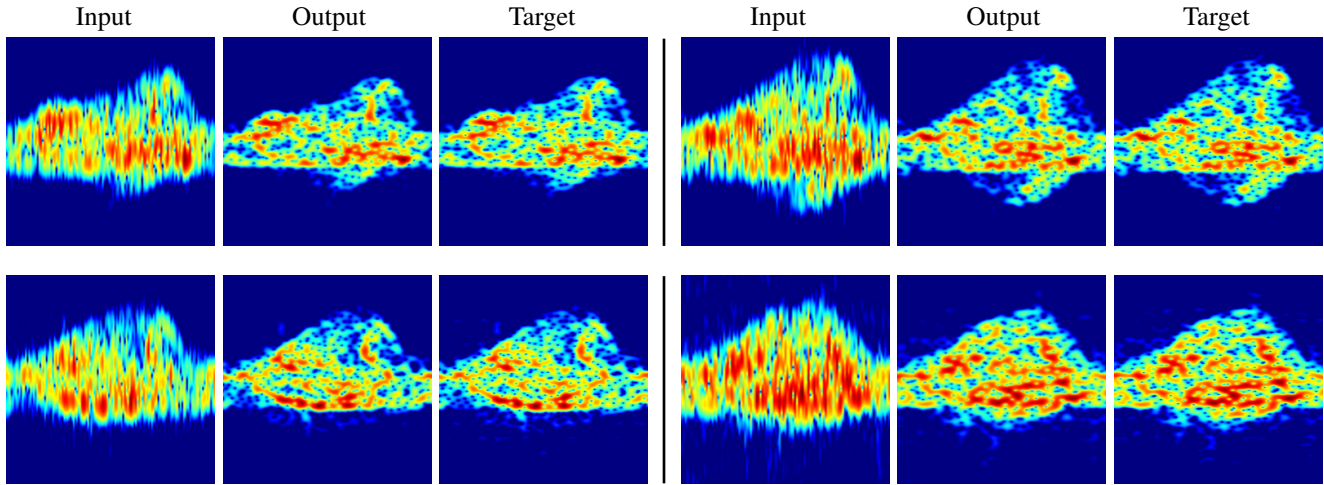


Fig. 3: Qualitative results for the super-resolution of human μ -D signature from a velocity resolution of 8 cm/s to 2 cm/s.

signature to the corresponding high-resolution counterpart, the network is able to generalize to never seen signatures for new subjects not included in the training dataset. Furthermore, the results are not based on simple interpolation of the pixel-wise values. They are capable of maintaining both the global consistency as well as the local structure of the signature. Such preserved vital information represents identifying fingerprints which is unique in the signature of every subject.

Since the results show a high visual similarity between the output and the target signature, a quantitative comparison between both images is applied across the used test set. The correlation between the super-resolved μ -D signatures and their corresponding high-resolution ground truths was calculated to be 0.979 according to the structural similarity index (SSIM) [27], and 0.953 according to the universal quality index (UQI) [28]. Moreover, a peak signal-to-noise ratio (PSNR) of 22.83 dB was achieved. This quantitative result indicates that minimal vital information loss occurred during the super-resolution process.

V. ULA USE-CASE AND DISCUSSION

In this section, the proposed super-resolution framework was applied for the enhancement of the range-azimuth measurements of a ULA radar. During the parametrization of such a radar system, the range and the 3dB-angular resolutions are given by

$$R_{res} = \frac{c}{2B} \quad \text{and} \quad \Psi_{res} = 0.89 \frac{\lambda}{Nd \cos(\theta)} \quad (5)$$

where B is the radar bandwidth, N is the number of receiving elements, $d = \frac{\lambda}{2}$ is the fixed antenna separation, θ is the angle of arrival. As such, a ULA radar experiences multiple design trade-offs. For instance, increasing the bandwidth B results in a better range resolution while having a detrimental effect on the maximum unambiguous range. Likewise, increasing the number of receiving elements enhances the radar angular resolution. However, it necessitates a longer acquisition time which may not be feasible in real-time applications [2], [3].

In this experiment, a dataset was collected over two separate acquisitions using a 77 GHz MIMO FMCW radar. This radar was operated effectively as a ULA with a single transmitting

element. The radar was oriented such that the azimuth beam steering is in the z -plane and three-steps of a stair case were considered as the objects of interest in the measurements. The experiment setup and the scanning orientation of the radar are illustrated in Fig. 4. The two radar acquisitions were parametrized with $B = 1.2$ GHz and $N = 4$ receiving elements for low-resolution acquisitions. For the high-resolution case, $B = 3.6$ GHz and $N = 8$ receiving elements. Thus, the range resolution is enhanced by a factor of three and the azimuth resolution by a factor of two. The proposed super-resolution framework was then trained to translate the low-resolution acquisitions into their corresponding high-resolution counterparts. A limited dataset of manually paired images were used for the training procedure.

The qualitative result of this preliminary experiment is shown in Fig. 4. It can be observed that lowering the utilized bandwidth and number of antennas deteriorates the resolution for both the range and azimuth such that the objects of interest can not be separated and are observed as a single entity. After the super-resolution translation by a cGAN, the range and azimuth resolutions are significantly enhanced and the three steps are clearly separated. The resulting range-azimuth spectrum looks almost the same as a radar with a parametrization of three-times the bandwidth and two-times the number of receiving elements.

To summarize, in this work we propose the use of a cGAN framework for the super-resolution of different radar spectra. This is achievable since the cGAN framework is trained using paired datasets of low and high-resolution acquisitions. This enables the framework to learn the joint distribution mapping low-resolution data to that of higher resolutions as well as grasp the underlying structures of radar spectra. However, our work is not without limitations. Obtaining such paired datasets is often challenging in many practical applications, e.g. for MIMO or range-Doppler measurements with extensive planning and acquisition effort. Moreover, the trained cGAN framework is sensitive to the choice of the radar system and parametrization used within the utilized training dataset. By using data from different setups and applications to train the model, a certain generalization capability could be achieved.

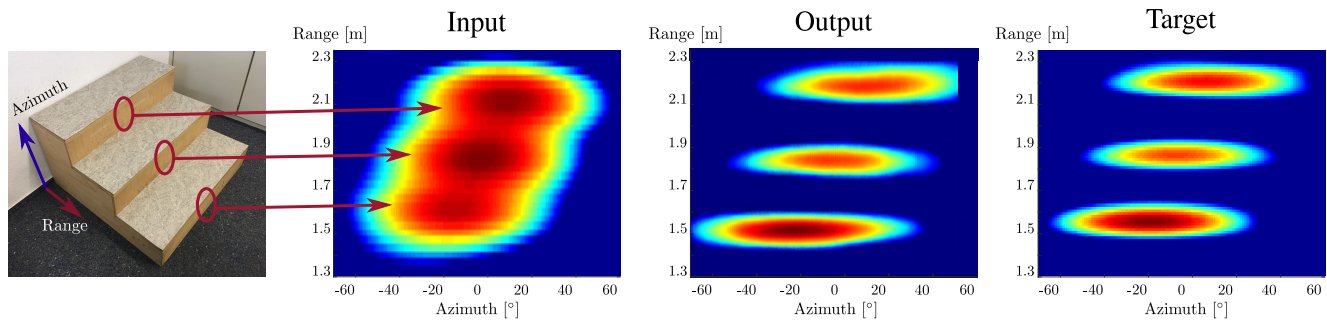


Fig. 4: Results for radar super-resolution in the range-azimuth plane.

VI. CONCLUSION

In this paper, a new technique for the super-resolution of radar signals was presented based on the utilization of a deep adversarial framework with additional non-adversarial losses. This technique aims to overcome the trade-offs in radar design parametrization by enhancing the post-processed 2D radar representations rather than operating on time-series IQ data.

To validate this proposition, experiments were carried out to enhance the velocity resolution, from 8 cm/s to 2 cm/s, for μ -D radar signatures while maximizing the unambiguous velocity range. Qualitative and quantitative analysis of the results indicate that the super-resolved μ -D signatures are highly correlated with their ground-truth high-resolution counterparts. Furthermore, the adversarial framework for super-resolution was applied in a preliminary experiment to enhance the resolution of range-azimuth measurements of a ULA radar with promising results.

In the future, we plan to examine the performance of the super-resolved μ -D signatures on post-processing tasks, such as object detection, localization and subject identification. We also aim to investigate the performance of adversarial super-resolution of ULA and MIMO radars more thoroughly by utilizing more test objects in training with a larger dataset. Additionally, extending the model to utilize recent unpaired adversarial translation techniques is crucial for extending the super-resolution framework for different radar measurements such as the range-Doppler maps in MIMO radars. Finally, in the future we plan to compare the deep learning-based super-resolution approach with other classical approaches based on IQ time-series.

REFERENCES

- [1] M. Murad *et al.*, "Requirements for next generation automotive radars," in *IEEE Radar Conference (RadarConf)*, April 2013, pp. 1–6.
- [2] M. A. Richards, *Fundamentals of radar signal processing*. New York: McGraw-Hill Education, 2005.
- [3] T. Geibig *et al.*, "Compact 3D imaging radar based on FMCW driven frequency-scanning antennas," in *IEEE Radar Conference (RadarConf)*, May 2016, pp. 1–5.
- [4] D. A. Dolgov and C. P. Urmson, "Modifying behavior of autonomous vehicles based on sensor blind spots and limitations," June 2016, uS Patent 9,367,065.
- [5] M. Ulrich *et al.*, "Person recognition based on micro-Doppler and thermal infrared camera fusion for firefighting," in *21st International Conference on Information Fusion (FUSION)*, July 2018, pp. 919–926.
- [6] R. Lasaponara and N. Masini, "Satellite synthetic aperture radar in archaeology and cultural landscape: An overview," *Archaeological Prospection*, vol. 20, no. 2, pp. 71–78, June 2013.
- [7] C. Ozdemir, *Inverse synthetic aperture radar imaging with MATLAB algorithms*. John Wiley & Sons, 2012, vol. 210.
- [8] M. G. Amin, *Through-the-wall radar imaging*. CRC press, 2017.
- [9] T. M. Carpenter *et al.*, "On the parametric and NEXRAD-radar sensitivities of a distributed hydrologic model suitable for operational use," *Journal of Hydrology*, vol. 253, no. 1-4, pp. 169–193, Nov. 2001.
- [10] S. Torres and C. Curtis, "Initial implementation of super-resolution data on the NEXRAD network," in *21st International Conference on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology*, 2007.
- [11] J. Hubbert *et al.*, "Data quality improvements using SZ phase coding," *32nd Conference on Radar Meteorology*, Jan. 2005.
- [12] G. Xu *et al.*, "Bayesian inverse synthetic aperture radar imaging," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 6, pp. 1150–1154, Nov. 2011.
- [13] S. Samadi *et al.*, "Sparse signal representation for complex-valued imaging," in *IEEE 13th Digital Signal Processing Workshop*, Jan. 2009, pp. 365–370.
- [14] Y. Xia *et al.*, "An improved MUSIC algorithm for angular super-resolution in scanning antenna," in *2013 International Conference on Computational Problem-Solving (ICCP)*, Oct. 2013, pp. 358–361.
- [15] I. Goodfellow *et al.*, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, 2014, pp. 2672–2680.
- [16] P. Isola *et al.*, "Image-to-Image translation with conditional adversarial networks," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017, pp. 5967–5976.
- [17] C. Ledig *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*, 2017, pp. 4681–4690.
- [18] K. Armanious *et al.*, "Retrospective correction of rigid and non-rigid MR motion artifacts using GANs," <https://arxiv.org/abs/1809.06276>, 2019, accepted to IEEE International Symposium for Biomedical Images (ISBI).
- [19] K. Armanious, Y. Mecky, S. Gatidis, and B. Yang, "Adversarial inpainting of medical image modalities," <https://arxiv.org/abs/1810.06621>, 2019, accepted to IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).
- [20] V. C. Chen, *The micro-Doppler effect in radar*. Artech House, 2011.
- [21] S. Abdulatif *et al.*, "Person identification and body mass index: A deep learning-based study on micro-Dopplers," <https://arxiv.org/abs/1811.07173>, 2019, accepted to IEEE Radar Conference (RadarConf).
- [22] S. Abdulatif, F. Aziz, B. Kleiner, and U. Schneider, "Real-time capable micro-Doppler signature decomposition of walking human limbs," in *IEEE Radar Conference (RadarConf)*, May 2017, pp. 1093–1098.
- [23] C. Wang *et al.*, "Perceptual adversarial networks for Image-to-Image transformation," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 4066–4079, Aug. 2018.
- [24] K. Armanious *et al.*, "MedGAN: Medical image translation using GANs," *CoRR*, vol. abs/1806.06397, 2018. [Online]. Available: <http://arxiv.org/abs/1806.06397>
- [25] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [26] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [27] Z. Wang *et al.*, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.
- [28] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430–444, Feb. 2006.