

Fully Convolutional Neural Networks for Automotive Radar Interference Mitigation

Nicolae-Catalin Ristea ¹, Andrei Anghel ¹, and Radu Tudor Ionescu ²

¹Affiliation not available

²University of Bucharest

October 30, 2023

Abstract

The interest of the automotive industry has progressively focused on subjects related to driver assistance systems as well as autonomous cars. In order to achieve remarkable results, cars combine a variety of sensors to perceive their surroundings robustly. Among them, radar sensors are indispensable because of their independence of light conditions and the possibility to directly measure velocity. However, radar interference is an issue that becomes prevalent with the increasing amount of radar systems in automotive scenarios. In this paper, we address this issue for frequency modulated continuous wave (FMCW) radars with fully convolutional neural networks (FCNs), a state-of-the-art deep learning technique. We propose two FCNs that take spectrograms of the beat signals as input, and provide the corresponding clean range profiles as output. We propose two architectures for interference mitigation which outperform the classical zeroing technique. Moreover, considering the lack of databases for this task, we release as open source a large scale data set that closely replicates real world automotive scenarios for single-interference cases, allowing others to compare objectively their future work in this domain. The data set is available for download at: <http://github.com/ristea/arim>.

Fully Convolutional Neural Networks for Automotive Radar Interference Mitigation

Nicolae-Cătălin Ristea

Department of Telecommunications
University Politehnica of Bucharest
Bucharest, Romania
r.catalin196@yahoo.ro

Andrei Anghel

Department of Telecommunications
University Politehnica of Bucharest
Bucharest, Romania
andrei.anghel@munde.pub.ro

Radu Tudor Ionescu

Department of Computer Science
University of Bucharest
Bucharest, Romania
raducu.ionescu@gmail.com

Abstract—The interest of the automotive industry has progressively focused on subjects related to driver assistance systems as well as autonomous cars. In order to achieve remarkable results, cars combine a variety of sensors to perceive their surroundings robustly. Among them, radar sensors are indispensable because of their independence of light conditions and the possibility to directly measure velocity. However, radar interference is an issue that becomes prevalent with the increasing amount of radar systems in automotive scenarios. In this paper, we address this issue for frequency modulated continuous wave (FMCW) radars with fully convolutional neural networks (FCNs), a state-of-the-art deep learning technique. We propose two FCNs that take spectrograms of the beat signals as input, and provide the corresponding clean range profiles as output. We propose two architectures for interference mitigation which outperform the classical zeroing technique. Moreover, considering the lack of databases for this task, we release as open source a large scale data set that closely replicates real world automotive scenarios for single-interference cases, allowing others to compare objectively their future work in this domain. The data set is available for download at: <http://github.com/ristea/arim>.

Index Terms—autonomous driving, automotive radar, interference mitigation, denoising, convolutional neural networks, deep learning.

I. INTRODUCTION

Nowadays, the automotive radar sensors are essential elements of driving assistance systems as well as of autonomous driving applications. Their main goal is to estimate the distance and velocity of objects on the road. However, the technical requirements increased steadily from simple detection to braking functions and smart environment perception tasks for self-driving cars. The most common radar sensors used in the automotive industry are frequency modulated continuous wave (FMCW) / chirp sequence (CS) radars, which transmit sequences of linear chirp signals. Furthermore, the amount of automotive radar systems on the streets is regularly increasing [1], and the high number of automotive radar systems on roads lead to a higher probability of interference between radar sensors.

The author wish to acknowledge all national funding authorities and the ECSEL Joint Undertaking, which funded the PRYSTINE project under the grant agreement 783190. The real data used for experimental validation were provided by NXP under the PRYSTINE cooperation agreement. Additionally, this work was co-funded through the Operational Program Competitiveness 2014-2020, Axis 1, contract no. 3/1.1.3H/24.04.2019, MySMIS ID: 121784.

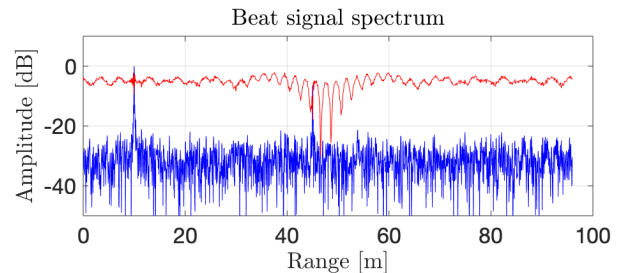


Fig. 1: Range profile magnitude of an FMCW radar sensor. The useful profile is shown in blue, while the profile with interference is shown in red.

Radio frequency interference (RFI) is a relevant issue for radar sensors as it could increase the effective noise floor, reduce sensitivity or create false detections [2]. We illustrate the negative effects of RFI in Figure 1, where the noise floor has risen with approximately 15 dB and the targets become almost undetectable. Therefore, the interference mitigation task is a crucial part of current and future radar sensors used in a traffic safety context.

In this paper, we propose a novel approach for radar interference mitigation that is based on fully convolutional neural networks (FCNs) [3]. This architecture is able to learn different types of patterns with a relatively small amount of learnable parameters, a fact that recommends them for real-time applications, such as autonomous driving. Our proposed architecture is fed with spectrograms of beat signals with noise and interference. The output of our neural network is the range profile magnitude with mitigated interference and noise. Additionally, because of the lack of public databases, we propose a novel synthetic database, with signals affected by an interference source, for the community to have a common ground for the evaluation and comparison of future methods.

In summary, our contribution is twofold:

- We propose a novel approach, based on FCN architectures for interference mitigation, which transforms a spectrogram into a range profile.
- We propose a radar interference data set with a wide range of signal parameters variations to be used as benchmark in future research.

II. RELATED WORK

Classical methods. State-of-the-art interference detection (mitigation) methods are typically classified according to the domain in which the interference is mitigated [4]–[9]: polarization, time, frequency, code and space. Polarization-based methods assume the use of cross-polarized antennas between the two interfering radars and the mitigation margin is around 20 dB, but reflections on the ground or other surrounding targets can severely reduce this margin. Time domain methods include the following approaches: use of a low transmit duty cycle (to reduce the probability of hitting other receivers), use of a short receive window (to reduce the probability of being hit by an interferer), or employing a variable pause between chirps or a variable chirp slope to avoid periodic interferences. Frequency domain methods imply a division of the authorized operating bandwidth into several sub-bands, such that nearby radars operate in different sub-bands. RFI mitigation in the coding domain involves the modulation of the radar waveforms with a device-specific code (to minimize cross-talk between radars, the codes of different devices should be orthogonal), whereas in the case of space domain techniques the antenna radiation pattern is adaptively configured to avoid interfering signals.

A particular class of methods are the strategic RFI mitigation techniques [4], which involve additional hardware and/or software, and rely on some of the basic ones. The classical strategic approaches are: “communicate and avoid” (requires inter-vehicle communication to avoid simultaneous transmission), “detect and avoid” (e.g., detects the interference in a sub-band and changes the operating sub-band of the radar), “detect and repair” (after detection, the measurement with interference is reconstructed), “detect and omit” (after detection, the measurements affected by interference are removed) and “listen before talk” (the radar transmits only when no other device is detected).

Different from all these methods, which rely on algorithms handcrafted by researchers, we propose an approach based on end-to-end learning from data. In order to obtain our approach, we constructed a realistic data set to learn deep neural networks.

Deep learning methods. Deep learning techniques have been applied in a wide range of tasks with remarkable results [10], [11]. One such task is image denoising, where deep learning achieved state-of-the-art results [12], outperforming classical approaches such as median or bilateral filtering. By transforming the radio signal into a spectrogram, the task of interference mitigation becomes similar to a task of image denoising. In this context, we propose to apply fully convolutional networks, a deep learning technique, to transform a noisy spectrogram into a clean range profile of a radar sensor. To our knowledge, there are only a handful of related works [13]–[15] that employ deep learning models for radar interference mitigation, but they have different architectures that are based on input and output pairs from the same domain, e.g. both their input and output are spectrograms [13]. In [13], the authors proposed a

convolutional neural network (CNN) to address RFI, aiming to reduce the noise floor while preserving the signal components of detected targets. The CNN architecture can be trained using either range processed data or range-Doppler (RD) spectra as inputs. The authors reported promising results, but they still have concerns regarding the generalization capacity on real data. Another approach that relies on CNNs is proposed in [16]. The authors employ an autoencoder based on the U-Net architecture [17], which performs interference mitigation as a denoising task directly on the range-Doppler spectrum. They surpass classical approaches, but the phase information cannot be fully preserved. Similarly, in [14], the network architecture is build upon CNNs, but residual connections, inspired from [18], were added. A different method is proposed in [15], which is based on applying a recurrent neural network model with Gated Recurrent Units [19] on the time domain signal to mitigate the interference. The authors reported better performance compared to existing signal processing methods and lower processing times.

III. METHOD

Radar signal model. In FMCW radar solutions, the transmitted signal $s_{TX}(t)$ is a chirp sequence, whose frequency usually follows a sawtooth pattern. The analytical signal $s_{TX}(t)$ in a sweep interval is defined as follows:

$$s_{TX}(t) = \begin{cases} e^{j2\pi(f_0 t + \frac{\alpha}{2} t^2)}, & \text{if } 0 \leq t \leq T_c \\ 0, & \text{otherwise} \end{cases}, \quad (1)$$

where t is the time domain variable, f_0 is the frequency at the initial moment $t = 0$, T_c is the frequency variation time interval (sweep time) and α is called the chirp rate calculated as $\alpha = \frac{B}{T_c}$, where B is the bandwidth.

The receive antenna collects the reflected signal $s_{RX}(t)$, which, for a single target, is defined as follows:

$$s_{RX}(t) = \underline{A} \cdot s_{TX}(t - \tau), \quad (2)$$

where τ is the propagation delay and $\underline{A} = A \cdot e^{j\phi}$ is the complex amplitude of the received signal.

The reflected signal is further mixed with the transmitted signal and low-pass filtered, resulting the beat signal $s_b(t)$, which is analytically computed as:

$$s_b(t) = s_{TX}(t) \cdot s_{RX}^*(t), \quad (3)$$

where $s_{RX}^*(t)$ represents the complex conjugate of $s_{RX}(t)$.

Interference signal model. In the presence of mutual interference, the radar transmits a signal which is reflected by a target and the receiving antenna collects a mix from two signals, the reflected and the interference signal. Consequently, the received signal is defined as follows:

$$s_{RX}(t) = \sum_{j=0}^{N_t} \underline{A}_j \cdot s_{TX}(t - \tau_j) + \sum_{k=0}^{N_{int}} s_{RFI,k}(t), \quad (4)$$

where N_t is the number of targets, N_{int} is the number of interferers and $\underline{A}$, τ are the corresponding parameters from Equation (2). An uncorrelated interfering signal $s_{RFI,k}(t)$ (with different modulation rate than the one of the transmitted chirp) translates, after mixing with the transmitted signal, into a chirp signal whose bandwidth is limited by the anti-aliasing

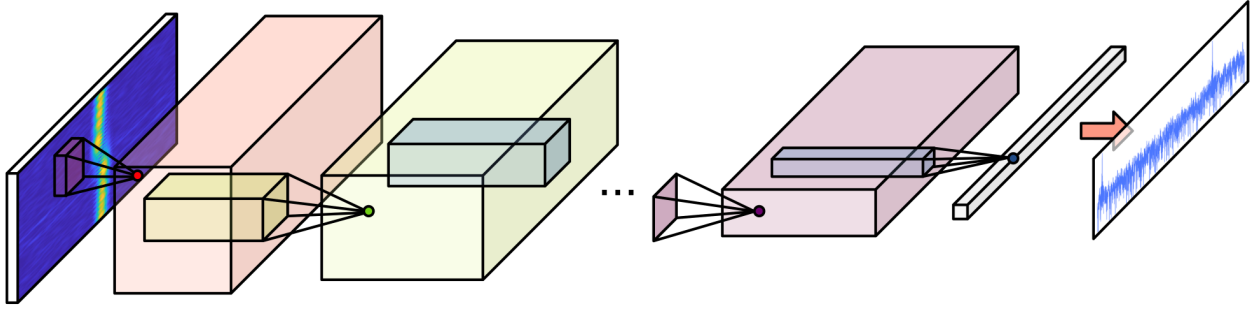


Fig. 2: The general architecture of our FCN models. The input spectrogram is processed through a series of conv blocks (composed of conv and pooling layers) until the vertical dimension is reduced to 1, while preserving the horizontal dimension. The output is a range profile without the interference removed by the FCN.

filter. Therefore an uncorrelated interference appears as a highly non-stationary component on the spectrogram, which is spread across multiple frequency bins. In the following, we limit the number of interferers to $N_{int} = 1$ and we consider only uncorrelated interfering sources.

Data preprocessing. Fully convolutional neural networks attain state-of-the-art results in computer vision, the convolutional operation being specifically designed for images. In order to apply FCNs to our tasks, we first need to transform the time domain signals into images. One of the most common approaches to obtain an image representation of a time domain signal is by computing a spectrogram using the discrete Short-Time Fourier Transform (STFT), as shown in the following equation:

$$STFT\{x[n]\}(m, k) = \sum_{n=-\infty}^{\infty} x[n] \cdot w[n - mR] e^{-j \frac{2\pi}{N_x} kn}, \quad (5)$$

where $x[n]$ is the discrete input signal, $w[n]$ is a window function, N_x is STFT length and R is the hop/step size [20]. There are a plethora of window functions proposed in literature, such as *hann*, *blackman* and others. We chose to perform the STFT with *hamming* window, which is described by the formula:

$$w[n] = 0.54 - 0.46 \cdot \cos\left(\frac{2\pi n}{N-1}\right), 0 \leq n \leq N-1, \quad (6)$$

where N is the window length.

Moreover, several time-frequency representations have been developed over time. The *Wigner-Ville* distribution is one of the techniques used to explore time-frequency information about a signal [21]. Wavelet analysis, through the *continuous wavelet transform* with different base functions, can also be used for this purpose. Nevertheless, in this study, we restrict ourselves to the spectrogram, and leave other time-frequency representations for future investigations.

Our goal is to obtain clean range profiles from beat signal spectrograms affected by noise and uncorrelated interference. We design our FCNs to provide the clean range profiles as output (during training, the FCNs have to learn to reproduce the ground-truth clean range profiles). For this reason, we perform a Fast Fourier Transform (FFT) of our time domain labels (to obtain the ground-truth clean range profiles) and train our networks to map the STFT input to the FFT output.

The FFT is computed as follows:

$$FFT\{x[n]\}(k) = \sum_{n=0}^{N_x-1} x[n] \cdot e^{-j \frac{2\pi}{N_x} kn}, \quad (7)$$

where N_x is the number of FFT points (same value as the number of STFT range bins).

Network architectures. We consider a beat signal with 1024 samples, and we design two FCN models for interference mitigation from single-channel spectrograms. The first FCN takes as input a spectrogram of 154×2048 components, while the second FCN takes as input a spectrogram of 1024×2048 components. The horizontal axis dimension (2048) corresponds to the number of FFT points in the range profile, while the vertical axis dimension (154 or 1024) is influenced by the computing step size of STFT. The dimensions of 154 or 1024 correspond to the steps 6 and 1, respectively. We design each FCN architecture to progressively reduce the dimension on the vertical axis to one component, while preserving the horizontal axis dimension. Therefore, both FCN models produce outputs of 1×2048 components that are interpreted as range profiles. In order to reach the same output size from different input sizes, the two FCN models have different depths. In Figure 2, we illustrate the generic architecture of both networks, which are exclusively composed of convolutional (conv) blocks.

To reduce the spectrogram of 154×2048 to a range profile of 1×2048 components, we propose a shallower FCN of 15 layers organized into 4 conv blocks. Each of the first 3 blocks are composed of 3 conv layers followed by a max-pooling layer, while the last block has 3 conv layers (without pooling). Each conv layer in the first block is formed of 8 filters. The number of conv filters doubles in each subsequent block. Based on this rule, the first two conv layers from the fourth block are formed of 64 filters. The last conv layer needs to squeeze the number of channels to one, hence it can have only one filter.

To reduce the spectrogram of 1024×2048 to a range profile of 1×2048 components, we propose a deeper FCN of 21 layers organized into 7 conv blocks. Each of the first 6 blocks are composed of 2 conv layers followed by a max-pooling layer, while the last block has 3 conv layers (without pooling). Each conv layer in the first block is formed of 8 filters. The number of conv filters doubles in each subsequent block, except for

TABLE I: Minimum and maximum values each parameter in our joint uniform distribution used for generating the samples in our database.

Parameter	Minimum	Maximum	Step
SNR [dB]	5	40	5
SIR [dB]	-5	40	5
Relative interference signal slope	0	1.5	0.1
Number of targets	1	4	-
Target amplitude	0.01	1	-
Target distance [m]	2	95	-
Target phase [rad]	$-\pi$	π	-

TABLE II: Fixed parameters for simulating a realistic radar sensor.

Parameter	Description	Value
B	Bandwidth	1.6 GHz
T_r	Time of chirp	25.6 μs
f_s	Sampling frequency	40 MHz
f_0	Radar central frequency	78 GHz

the second and the fourth blocks, which keep the number of filters from the previous blocks. As for the shallow FCN, the last conv layer has only one filter.

For both architectures, the conv filters with 5×5 spatial support are applied at stride 1. Zero padding is added to preserve the horizontal dimension of the activation maps. Except for the very last conv layer, all convolutional layers are followed by ReLU activations. The pooling filters are of size 2×1 , reducing the size of the activation maps on the vertical axis only. Zero padding for the max-pooling layers is added only when we need to make sure that the input activation maps have an even size. Both networks are trained using the Adam optimizer [22], using the mean squared error as loss function.

IV. DATA SET

It is well known that a large database is a key factor in the training process of deep neural networks with high generalization capacity. To the best of our knowledge, there are no public in-the-wild (or generated) databases for the interference mitigation task, mainly because of the difficulty imposed by the process of data acquisition and data labeling. In this context, it is hard to compare novel approaches with the previous ones in an objective manner.

In this paper, we propose a novel large scale database consisting of 48,000 samples, generated automatically while trying to replicate a realistic automotive scenario with one interference source. We generate every sample using specific randomly selected values for the set of parameters listed in Table I. While the values corresponding to some parameters are selected using a fixed step between the minimum and the maximum values specified in Table I, the values corresponding to the other parameters are randomly selected using an uniform distribution between the minimum and the maximum values. More precisely, we use linear variation with a fixed step for the signal to noise ratio (SNR), the signal to interference ratio (SIR) and the interference slope parameters. The number of targets as well as the distance, the amplitude and the phase

of each target are random variables that follow an uniform distribution. We added a random phase to each target to obtain more realistic radar signals.

Concerning the simulated radar sensor, we considered a fixed set of parameters such as bandwidth, sweep time, sampling frequency and central frequency. The exact values used for these parameters are listed in Table II.

Since we can control all factors during sample generation, we can produce an exact copy of each signal without interference. First of all, the clean copy can be used as ground-truth label when training a machine learning model. Second of all, it provides the means to conduct an objective assessment of the performance, by comparing the output predicted by the model with the corresponding ground-truth (expected) output. Consequently, a data sample is composed of:

- a time domain signal without interference;
- a time domain signal with interference;
- a label vector with complex amplitude values in target locations.

We randomly split our data samples into a training set of 40,000 samples and a test set of 8,000 samples. This will allow future works to directly compare novel results with previous results, without having to re-implement preceding methods in order to reproduce the corresponding results. Our data set is freely available for download at: <http://github.com/ristea/arim>.

V. EXPERIMENTS

Evaluation metrics. Usually the goal in radar signal processing is to maximize the detection performance. Thus, a rather intuitive measure is the area under the receiver operating characteristics curve (AUC), which describes the ability to disentangle targets from noise at various thresholds. Another performance indicator is the mean absolute error (MAE) in decibels (dB) between the range profile amplitude of targets computed from label signals and the amplitude of targets from predicted signals. In our evaluation, we employed the AUC, the MAE and the signal-to-noise ratio (SNR).

Baseline. A very common approach to eliminate the interference is to replace amplitudes higher than a specific threshold with zero (e.g. [4], [9]). This method is denoted as *zeroing*.

Hyperparameter tuning. We tune the hyperparameters of our FCN models on a validation set. We kept 20% of the training set (8,000 samples) for validation. We used the same hyperparameters for both architectures, in order to minimize the chance of overfitting in hyperparameter space. We trained the models for 100 epochs with a mini-batch size of 10 samples. We set the learning rate to 10^{-5} and we used a weight decay of 10^{-5} . These parameters are obtained using grid search. In a similar fashion, we employed grid search on the validation set to find the threshold parameter for our baseline, the zeroing method.

Results. We compared our FCN models with a zeroing benchmark and an oracle based on the ground-truth labels. We present the corresponding results in Table III. We included the oracle in order to show the achievable upper bound score for each metric, e.g. the maximum AUC score on both validation

TABLE III: Validation and test results provided by our shallow and deep FCN models versus an oracle based on true labels and a baseline based on zeroing. A higher SNR value is better, a higher AUC value is better and a lower MAE value is better. The best results (excluding the oracle) are highlighted in bold.

Method	Validation set			Test set		
	SNR	AUC	MAE (dB)	SNR	AUC	MAE (dB)
Oracle (true labels)	31.66	0.978	0	31.62	0.978	0
Zeroing	28.68	0.967	2.66	28.71	0.967	2.70
Shallow FCN	29.03	0.965	1.82	29.03	0.965	1.82
Deep FCN	31.61	0.972	0.91	30.99	0.970	0.89

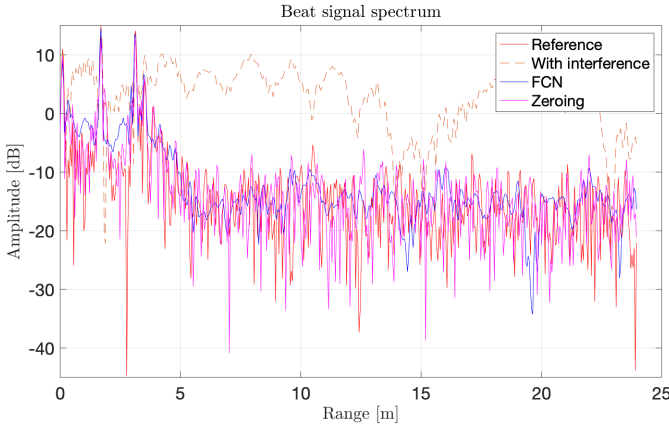


Fig. 3: Results for radar interference mitigation with our best FCN model. For comparison, we also added the ground-truth signal, the signal with interference and the signal with mitigated interference by the zeroing method.

and test is 0.978. Comparing our shallow FCN model with the zeroing baseline, we observe that our network attains better SNR and MAE scores, but it seems to underperform according to the AUC measure. Our deep FCN provides superior results for all three metrics, surpassing both the zeroing baseline and the shallow FCN. In terms of SNR and AUC, our deep FCN seems to attain performance levels quite close to the oracle, e.g. the AUC of our deep FCN is 0.008 under the AUC of the oracle on the test set.

In addition to the results on our data set, we assessed the generalization capacity of our deep FCN on real data, by testing it on samples provided by the NXP company, which were captured with the NXP TEF810X 77 GHz radar transceiver, from real-world scenarios. In Figure 3, we show that our model can perform interference mitigation on real radar signals, having an output very similar with the reference signal.

VI. CONCLUSION

In this paper, we introduced two fully convolutional networks for automotive radar interference mitigation and a large scale database of radar signals simulated in realistic settings. We compared our FCN models with the zeroing baseline in a comprehensive experiment, showing that our deep FCN provides superior results. We also released our novel data set to allow objective comparison in future work. To our

knowledge, we are the first to establish a benchmark data set for automotive radar interference mitigation. In future work, we aim to analyze the scenario with multiple interference sources.

REFERENCES

- [1] M. Kunert, “The EU project MOSARIM: A general overview of project objectives and conducted work,” in *Proceedings of EuRAD*, pp. 1–5, 2012.
- [2] G. M. Brooker, “Mutual interference of millimeter-wave radar systems,” *IEEE Transactions on Electromagnetic Compatibility*, vol. 49, no. 1, pp. 170–181, 2007.
- [3] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *Proceedings of CVPR*, pp. 3431–3440, 2015.
- [4] M. Kunert, F. Bodereau, M. Goppelt, C. Fischer, A. John, T. Wixforth, A. Ossowska, T. Schipper, and R. Pietsch, “D1.5 - Study on the state-of-the-art interference mitigation technique, MORE Safety for All by Radar Interference Mitigation (MOSARIM) project,” tech. rep., Robert Bosch GmbH, 2010.
- [5] F. Uysal, “Synchronous and asynchronous radar interference mitigation,” *IEEE Access*, vol. 7, pp. 5846–5852, 2019.
- [6] G. Kim, J. Mun, and J. Lee, “A Peer-to-Peer Interference Analysis for Automotive Chirp Sequence Radars,” *IEEE Transactions on Vehicular Technology*, vol. 67, pp. 8110–8117, Sep. 2018.
- [7] Z. Xu and Q. Shi, “Interference Mitigation for Automotive Radar Using Orthogonal Noise Waveforms,” *IEEE Geoscience and Remote Sensing Letters*, vol. 15, pp. 137–141, Jan 2018.
- [8] J. Bechter, M. Rameez, and C. Waldschmidt, “Analytical and experimental investigations on mitigation of interference in a dbf mimo radar,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, pp. 1727–1734, May 2017.
- [9] F. Laghezza, F. Jansen, and J. Overdevest, “Enhanced Interference Detection Method in Automotive FMCW Radar Systems,” in *2019 20th International Radar Symposium (IRS)*, pp. 1–7, June 2019.
- [10] P. Soviany and R. T. Ionescu, “Optimizing the trade-off between single-stage and two-stage deep object detectors using image difficulty prediction,” in *Proceedings of SYNASC*, pp. 209–214, 2018.
- [11] M.-I. Georgescu, R. T. Ionescu, and N. Verga, “Convolutional Neural Networks with Intermediate Loss for 3D Super-Resolution of CT and MRI Scans,” *arXiv preprint arXiv:2001.01330*, 2020.
- [12] P. A. Bricman and R. T. Ionescu, “CocoNet: A deep neural network for mapping pixel coordinates to color values,” in *Proceedings of ICONIP*, pp. 64–76, 2018.
- [13] J. Rock, M. Toth, E. Messner, P. Meissner, and F. Pernkopf, “Complex Signal Denoising and Interference Mitigation for Automotive Radar Using Convolutional Neural Networks,” *arXiv preprint arXiv:1906.10044*, 2019.
- [14] W. Fan, F. Zhou, M. Tao, X. Bai, P. Rong, S. Yang, and T. Tian, “Interference Mitigation for Synthetic Aperture Radar Based on Deep Residual Network,” *Remote Sensing*, vol. 11, no. 14, p. 1654, 2019.
- [15] J. Mun, H. Kim, and J. Lee, “A Deep Learning Approach for Automotive Radar Interference Mitigation,” in *Proceedings of VTC-Fall*, pp. 1–5, 2018.
- [16] J. Fuchs, A. Dubey, M. Lübke, R. Weigel, and F. Lurz, “Automotive Radar Interference Mitigation using a Convolutional Autoencoder,” in *Proceedings of RADAR*, 04 2020.
- [17] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Proceedings of MICCAI*, pp. 234–241, Springer, 2015.
- [18] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of CVPR*, pp. 770–778, 2016.
- [19] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” *arXiv preprint arXiv:1412.3555*, 2014.
- [20] J. B. Allen and L. R. Rabiner, “A unified approach to short-time Fourier analysis and synthesis,” *Proceedings of the IEEE*, vol. 65, pp. 1558–1564, Nov 1977.
- [21] L. Cohen, “Time-frequency distributions—a review,” *Proceedings of the IEEE*, vol. 77, no. 7, pp. 941–981, 1989.
- [22] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in *Proceedings of ICLR*, 2015.