Deep Learning-Based Radar Target Discrimination in Meteorological Clutter: A Preliminary Study

Guilherme Moura Delfino[1], Neuton Severo de Farias Neto[1], Leandro Guimarães Figueroa Pralon[2], Renato Machado[1], Dimas Irion Alves[1], Sarah Negreiros de Carvalho Leite[1]

[1]Telecommunications Department, Electronics Engineering Division, Aeronautics Institute of Technology, São José dos Campos, SP, Brazil

[2]Brazilian Army Technology Center, Rio de Janeiro, RJ, Brazil

***Abstract*—Radar target detection in meteorological clutter remains difficult in surveillance systems. Under adverse weather, conventional detection algorithms may generate excessive false alarms and degrade target tracking. This work presents a preliminary study on the use of deep learning for radar target discrimination using complex Doppler signatures extracted from real operational measurements of the SABER M200 Vigilante radar. The proposed approach organizes radar samples as structured tensors of Doppler signatures from multiple pulse repetition frequencies (PRFs) and receiver channels, preserving the real and imaginary components of the complex signal to allow the network to exploit both amplitude and phase information. A compact hybrid 2D/1D residual convolutional neural network ($\approx$147,000 parameters) is trained on a strict temporal partition and evaluated against a sealed cross-day dataset. The proposed model achieves an F_2 score of 0.910 and ROC-AUC of 0.999 on the cross-day test, outperforming linear baselines by more than 0.75 in F_2 and demonstrating feasibility on operational data.**

***Index Terms*—Radar signal processing, deep learning, meteorological clutter, Doppler processing, convolutional neural networks**

I. Introduction

Surveillance radar systems play a fundamental role in airspace monitoring and target tracking. However, detecting aerial targets under adverse meteorological conditions is challenging because meteorological clutter degrades conventional detection algorithms, increasing the Doppler background level and generating excessive false alarms.

Conventional techniques rely on threshold-based approaches such as Constant False Alarm Rate (CFAR) detectors [1]. Although widely deployed, these methods degrade in non-homogeneous environments where clutter statistics vary spatially and temporally, making discrimination between clutter and targets particularly difficult when both occupy similar regions of the Range-Doppler domain.

Polarimetric diversity is an established approach for discriminating scatterer types in weather-radar systems [2]. Spectral-polarimetric methods exploit depolarization to separate precipitation from ground clutter [3], [4]. However, these require dual- or full-polarimetric acquisition, which is unavailable in the single-polarization S-band radar considered here. This motivates discrimination techniques that rely solely on the complex Doppler, multi-PRF, and multi-receiver structure present in conventional surveillance measurements.

Deep learning has emerged as a powerful tool for radar signal analysis [5]. Convolutional Neural Networks (CNNs) in particular learn discriminative representations directly from structured radar data without relying on analytical clutter models [6], [7], exploiting spatial, spectral, and temporal correlations that are difficult to model with conventional signal processing [8], [9].

This work investigates the feasibility of CNN-based target discrimination under meteorological clutter using complex Doppler signatures from the SABER M200 Vigilante radar. The central hypothesis is that the joint observation of Doppler responses across multiple PRFs and receiver channels preserves discriminative characteristics that CNNs can exploit, including under a strict cross-day generalization test in which no information from the held-out day informed any modelling decision.

II. Database Generation

The dataset was collected during a real operational field experiment in Campinas, Brazil, using the SABER M200 Vigilante radar, a ground-based S-band phased-array pulse-Doppler surveillance radar developed through a cooperation between the Brazilian Army and Embraer [10]. The radar remained stationary while monitoring real aerial traffic. Meteorological conditions produced localized clutter that affected portions of the coverage volume.

A. Signal Processing and Data Extraction

The radar signal processing chain follows a conventional pulse-Doppler architecture. After coherent acquisition and range-Doppler processing, detection candidates are identified through CFAR processing on Range-Doppler maps [1]. The complex Doppler signature for range bins with significant Doppler activity is preserved for offline analysis. The proposed network does not receive binary CFAR decisions. It receives the full I/Q Doppler response, preserving spectral structure, Doppler distribution, and phase characteristics. The simplified extraction pipeline is shown in Fig. 1.

G. M. Delfino, gmdelfino@hotmail.com; N. S. de Farias Neto, neutonsevero@gmail.com; L. G. F. Pralon, pralon.leandro@eb.mil.br; R. Machado, rmachado@ita.br; D. I. Alves, dimasirion@ita.br; S. N. de Carvalho Leite, sarahnc@ita.br.

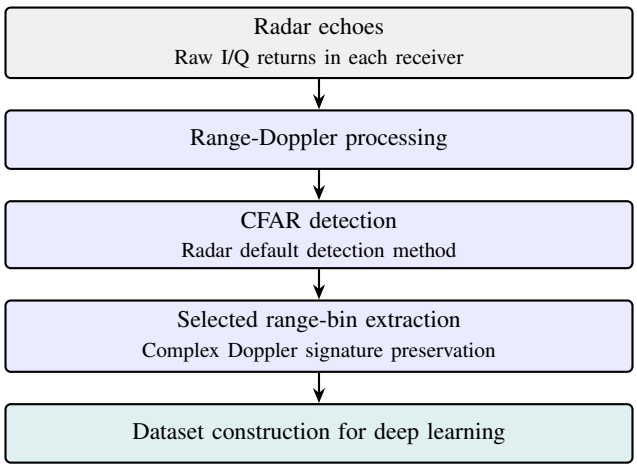

Fig. 1. Simplified radar signal processing and dataset extraction pipeline. Range bins presenting significant Doppler activity after CFAR-based processing are selected and their complex Doppler signatures are preserved for deep learning-based analysis.

B. Annotation and Temporal Partitioning

Target samples were manually annotated from range–time visualizations, with annotations corroborated by secondary surveillance radar (SSR) data. Target-related samples show coherent trajectories, and clutter samples show spatially extended patterns. Labels should be understood as practical experimental labels, not perfect ground truth.

A two-stage sanitization procedure was applied to remove duplicate and conflicting samples. First, intra-class duplicates were removed using the joint range-bin and timestamp key, eliminating repeated detections of the same range-time cell. Second, samples present in both classes with the same range-bin and timestamp were removed from both, discarding instances labelled inconsistently as target and clutter.

Data were partitioned by scan order to prevent temporal leakage. Cells sharing an antenna scan are correlated, so no scan straddles a partition boundary. C1 comprises two temporal acquisition segments collected on the same day. Segment 1 is assigned entirely to validation. Within Segment 0, a contiguous interior block is also assigned to validation. Two-scan guard intervals flanking the block are discarded. All remaining Segment 0 scans form the training set. A dataset acquired on a different day (C2) was held out as an external test set and was not used during model development. The resulting distribution is shown in Table I and the range-time view with partition assignments in Fig. 2.

TABLE I
DATASET DISTRIBUTION BY SUBSET AND CLASS, AFTER TWO-STAGE SANITIZATION. C1 IS SPLIT BY SCAN ORDER; C2, ACQUIRED ON A DIFFERENT DAY, IS THE HELD-OUT EXTERNAL TEST SET.

Subset	Clutter	Target	Total
C1 training	66,636	3,535	70,171
C1 validation	50,248	1,901	52,149
C2 ext. test	23,895	637	24,532

III. Radar Signal Representation

Each sample is the complex Doppler-domain signature for a selected range bin, acquired across multiple PRF bursts and receiver channels [11]. The measurement arrives as a one-dimensional complex vector. In this vector the Doppler index varies fastest, followed by the receiver channel, then the PRF burst. After partitioning into PRF blocks and demultiplexing per receiver, the signature is reorganised into a tensor:

$$X \in \mathbb{C}^{C \times P \times D},$$

where C is the number of receivers, P the PRF bursts, and D the Doppler bins. Due to confidentiality constraints, the exact values of C, P and D are not disclosed, but the physical structure is preserved in full.

A. Real and Imaginary Channel Decomposition

Because the proposed architecture uses real-valued convolutions, X is converted to a real-valued representation by concatenating the real and imaginary parts along the channel axis:

$$\tilde{X} \in \mathbb{R}^{2C \times P \times D}.$$

This preserves all information in the original complex signal while remaining compatible with standard CNN operations. Unlike magnitude-only representations, this retains phase information. Section V confirms that phase information is critical: a magnitude-only model (A6) degrades by 0.12 in F_2.

IV. Proposed Neural Network Architecture

We propose a compact hybrid two-stage residual CNN operating directly on $\tilde{X}$. The network has five macro blocks: a front end, a 2D PRF-Doppler residual stage, a transition, a 1D Doppler residual stage, and a decision head.

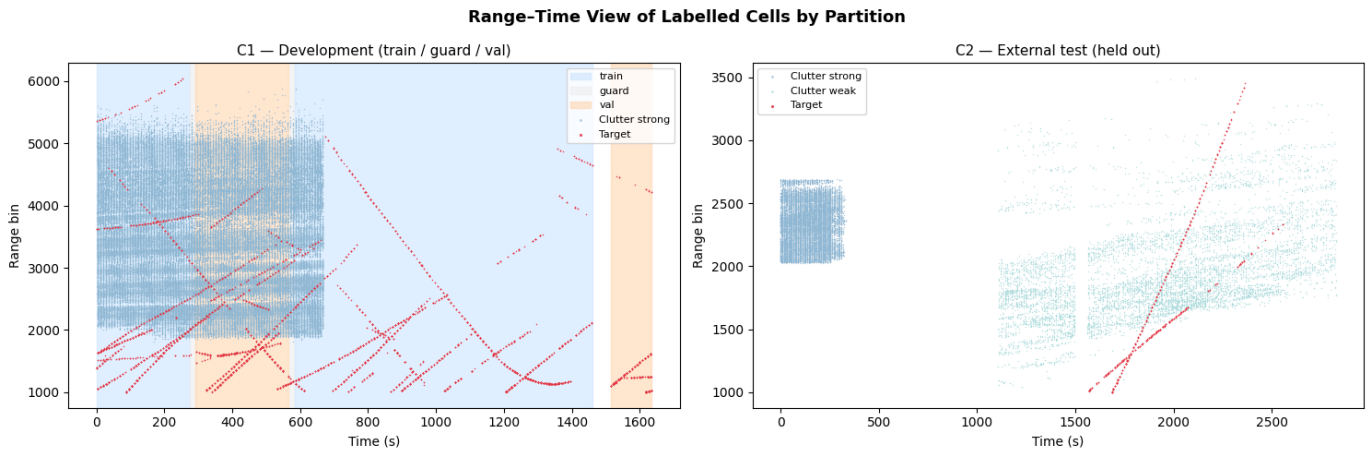

Fig. 2. Range-Time view of all labelled cells. *Left*: C1 development set with temporal partition highlighted (train/guard/val). Target trajectories appear as diagonal lines (red dots), clearly separable from meteorological clutter (blue cloud). *Right*: C2 external test set, containing strong and weak clutter alongside two clearly visible target trajectories.

A. Input Representation

The network is applied to the real-valued tensor $\tilde{X} \in \mathbb{R}^{2C\times P\times D}$ defined in Section III. I/Q components and receivers act as the $2C$ channels, while the PRF and Doppler axes form the spatial grid on which the 2D convolutions act. A single 2D kernel therefore spans several PRF bursts and Doppler bins at once, so the early layers learn how a return evolves jointly across slow time and Doppler rather than along either axis alone.

B. Macro Structure

The front end applies input BatchNorm directly to the $2C$-channel Re/Im tensor. Input normalization keeps the network self-contained and facilitates its integration into a radar processing chain. The small configuration uses no stem, so the first 2D block operates on the raw input channels. The 2D PRF-Doppler stage has two residual downsampling blocks that halve both the PRF and Doppler axes, progressively fusing the PRF dimension to a single row. A transition squeezes this axis into a Doppler sequence. Two 1D Doppler residual blocks refine the sequence. The decision head concatenates global average and max pooling and outputs a target-presence logit.

The two-stage design matches the structure of the signal. While both PRF and Doppler axes carry information, the 2D stage first learns their joint structure and fuses the PRF dimension to a single row. Once that axis is consumed, the remaining variation lies along Doppler alone, so 1D convolutions suffice and avoid the cost of 2D kernels over a degenerate axis. This keeps the network compact while still exploiting cross-PRF correlations where they are present.

Every block follows the residual mapping

$$\mathbf{Y} = \mathrm{ReLU}\big(\mathcal{F}(\mathbf{X}) + \mathcal{S}(\mathbf{X})\big), \tag{1}$$

where $\mathcal{F}$ is the main path and $\mathcal{S}$ the shortcut. Within this mapping, the 2D blocks downsample both axes with even kernels and strides of two, using asymmetric padding to preserve axis compatibility, while the 1D blocks use same-length padding throughout. The complete network configuration is summarized in Table II. The model comprises 147,377 trainable parameters.

TABLE II
PROPOSED HYBRID 2D/1D RESIDUAL CNN.

Block	Configuration	Output
Front end		
Input	tensor	$[2C, P, D]$
BN2D	input normalization	$[2C, P, D]$
2D PRF-Doppler stage (no stem; starts from $2C$ channels)		
Res-2D-1	$[3\times9] + [2\times4, \mathrm{s2}] \oplus [1\times1, \mathrm{s2}]$, $2C\to32$	$[32, P/2, D/2]$
Res-2D-2	$[1\times9] + [2\times4, \mathrm{s2}] \oplus [1\times1, \mathrm{s2}]$, $32\to64$	$[64, 1, D/4]$
Squeeze PRF	remove singleton	$[64, D/4]$
1D Doppler stage		
Res-1D-1	$[7] + [5] \oplus \mathrm{id}$	$[64, D/4]$
Res-1D-2	$[5] + [3] \oplus [1]$	$[32, D/4]$
Decision head		
Pool	GAP+GMP, concat	[64]
Output	Dropout, Linear	[1]

"+" chains main-path convolutions. "$\oplus$" is residual addition with shortcut + ReLU (Eq. 1). Strides are 1 unless noted.

C. Loss and Training

The model was trained using binary cross-entropy with logits:

$$\mathcal{L} = -\frac{1}{N}\sum_{n=1}^{N}\big[y_n \log\sigma(z_n) + (1-y_n)\log(1-\sigma(z_n))\big], \tag{2}$$

where σ is the logistic sigmoid and $y_n \in \{0,1\}$ the label. Class imbalance ($\approx 19{:}1$ clutter to target in training) is handled by weighting the positive term by $w_+ = 19$.

Hyperparameters were optimized through a sequential coordinate search on the C1 validation set using 30-epoch

training trials. The selected configuration employed the Adam optimizer with a cosine annealing learning rate schedule (maximum learning rate of 3×10^{-3}), ReLU activations, and 2D spatial dropout (dropout rate of 0.1) applied to the feature maps. The final model was then trained on the C1 training set for up to 50 epochs. The checkpoint achieving the highest validation F_2 score was selected at epoch 25.

D. Evaluation Metric

Standard accuracy is dominated by the majority clutter class ($\approx 19{:}1$). Therefore, the F_β score with $\beta = 2$ was adopted as the primary evaluation metric:

$$F_2 = \frac{5\,P \cdot R}{4\,P + R}, \tag{3}$$

where P and R denote precision and recall, respectively. This metric places greater emphasis on recall than precision. In surveillance applications, missing a target (i.e., reducing the probability of detection, P_d) is operationally more costly than generating a false alarm. Consequently, the F_2 score better reflects the operational priorities of the target discrimination task. All model selection decisions were based exclusively on the C1 validation F_2. The C2 test set was not used during model development

V. Results and Discussion

A. Baselines

Two linear classifiers were trained on the same temporal splits. A logistic regression (LR) model on magnitude features achieves $F_2 = 0.635$ on C1-val but collapses to $F_2 = 0.153$ on C2, showing that the learned representation does not generalise across days. An SVM on the raw complex signal performs worse in-distribution ($F_2 = 0.167$) and similarly on C2 ($F_2 = 0.117$). These results confirm that the problem is not linearly separable and motivate the architecture in Section IV.

B. Internal Performance (C1 Validation)

At the operating threshold of 0.655 selected on C1-val, the baseline CNN (A1) achieves $F_2 = 0.991$ and PR-AUC $= 0.995$. Table III shows the confusion matrix: of 1,901 target samples only 7 are missed, and of 50,248 clutter samples only 58 produce false alarms.

TABLE III
Confusion matrix on C1 validation (threshold 0.655): 99.6% target detection and 99.9% clutter rejection ($F_2 = 0.991$).

		Predicted	
		Target	Clutter
Actual	Target	1,894	7
	Clutter	58	50,190

C. Cross-Day Generalization (C2)

Table IV compares all models on the sealed C2 test. The threshold is fixed at 0.655 from C1 val with no re-tuning. Fig. 3 shows the ROC and precision-recall curves on C2 for the CNN baseline and both linear models.

TABLE IV
Cross-day generalization on the sealed C2 test set, threshold fixed at 0.655 (no re-tuning).

Model	F_2	AUC-ROC
LR (magnitude)	0.153	0.747
SVM (complex)	0.117	0.522
CNN Re/Im (A1)	0.910	0.999

The CNN (A1) achieves $F_2 = 0.910$ on C2, outperforming the best linear baseline by $+0.757$. At this operating point the CNN detects $P_d = 100\%$ of targets at $P_{fa} = 10\%$ and $P_d = 99.7\%$ at $P_{fa} = 1\%$, thresholds at which both linear baselines detect fewer than 2% of targets. The fixed threshold of 0.655 is one point on these curves, not the only usable one. Because the full detection and false-alarm trade-off is traced by sweeping the threshold, the operating point can be moved to match operational priorities. Lowering it raises P_d at the cost of higher P_{fa}, while raising it does the reverse. The C2 ROC curve in Fig. 3 stays close to the top-left corner (AUC-ROC 0.999), so a wide range of thresholds yields high detection at low false alarm. This margin is the practical value of the model: a deployment can select its own P_d or P_{fa} target without retraining, and the threshold transferred unchanged from C1 to C2, which indicates the operating point is stable across days.

D. Ablation Study

Table V ranks all ablations by C2 F_2. Each configuration modifies one design choice of the baseline CNN (A1).

TABLE V
Ablation results sorted by C2 F_2. Each variant modifies one factor from the baseline (A1).

ID	Description	C1-val F_2	C2 F_2
A1	Baseline Re/Im	0.991	0.910
A7	Magnitude + phase	0.979	0.877
A3	Per-sample normalization	0.977	0.871
A6	Magnitude only	0.980	0.794
A4	Phase-rotation aug.	0.981	0.792
A7b	Magnitude + unwr. ph.	0.982	0.787
A0	LR baseline (magnitude)	0.635	0.153

The ablation isolates phase information as the decisive factor: switching to magnitude-only (A6) reduces C2 F_2 to 0.794 ($\Delta = 0.116$ below A1), confirming that the Re/Im channel decomposition provides complementary information rather than just more features. Variants that combine magnitude with explicit phase channels (A7, A7b) also underperform the joint

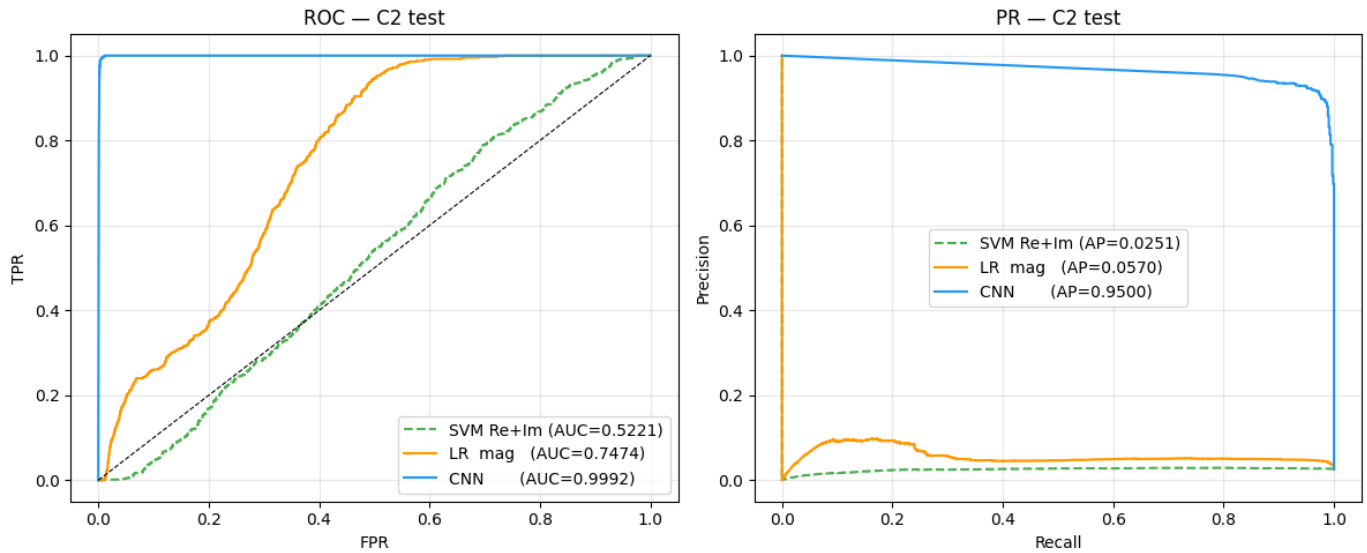

Fig. 3. ROC (left) and Precision-Recall (right) curves on the C2 cross-day test set for the CNN baseline (A1), logistic regression (LR), and SVM. The CNN dominates across all operating points (AUC-ROC 0.999, AP 0.950).

Re/Im representation, indicating that the benefit comes from how real and imaginary parts are presented to the convolutions rather than from the mere availability of phase. Phase-rotation augmentation (A4) and per-sample normalization (A3) both reduce cross-day F_2, suggesting that modifying the phase structure or removing global amplitude information is harmful.

VI. Conclusion

We presented a compact hybrid 2D/1D residual CNN (147,377 parameters) for binary target-presence detection under rain clutter, operating directly on the complex multi-PRF, multi-receiver Doppler tensor produced by the SABER M200 Vigilante radar. The Re/Im channel decomposition is essential to cross-day performance. On C1 validation the network achieves $F_2 = 0.991$ with only 7 missed targets and 58 false alarms among 52,149 samples. On the sealed cross-day C2 set, the network attains $F_2 = 0.910$ and AUC-ROC $= 0.999$, outperforming the best linear baseline by 0.757 in F_2. The CNN detects 99.7% of C2 targets at $P_{fa} = 1\%$, thresholds at which linear baselines detect fewer than 2%.

These results confirm the central hypothesis of this work: the joint multi-PRF, multi-receiver complex Doppler representation contains discriminative information that can be effectively exploited by a compact CNN. This demonstrates the feasibility of deep learning-based target discrimination in meteorological clutter using real operational radar data.

A. Limitations

Results come from a single operational campaign under one meteorological event type and should be interpreted as preliminary feasibility evidence. Annotation from Range-Time visualizations may contain residual uncertainty for weak targets in clutter-dominated regions. Performance may also depend on campaign-specific preprocessing and clutter characteristics, which motivates evaluation across more diverse conditions.

B. Future Work

Directions include evaluation across multiple campaigns and event types, complex-valued architectures [12] that natively process I/Q, uncertainty quantification for adaptive threshold selection, domain adaptation for cross-day shift, and multi-class discrimination between clutter types.

References

[1] H. Rohling, "Radar CFAR thresholding in clutter and multiple target situations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-19, no. 4, pp. 608–621, 1983.

[2] V. N. Bringi and V. Chandrasekar, *Polarimetric Doppler Weather Radar: Principles and Applications*. Cambridge University Press, 2001.

[3] M. He, Y. Nian, X. Wang, S. Xiao, and Y. Li, "Improved clutter suppression algorithm for atmospheric target detection using polarimetric doppler radar," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 4, pp. 911–922, 2011.

[4] J. Yin, C. Unal, and H. Russchenberg, "Narrow-band clutter mitigation in spectral polarimetric weather radar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 9, pp. 4655–4667, 2017.

[5] Z. Geng, H. Yan, J. Zhang, and D. Zhu, "Deep-learning for radar: A survey," *IEEE Access*, vol. 9, pp. 141 800–141 818, 2021. [Online]. Available: https://doi.org/10.1109/ACCESS.2021.3119561

[6] D. Cha, S. Jeong, M. Yoo, J. Oh, and D. Han, "Multi-input deep learning based fmcw radar signal classification," *Electronics*, vol. 10, no. 10, p. 1144, 2021. [Online]. Available: https://doi.org/10.3390/electronics10101144

[7] S. Capobianco, L. Facheris, F. Cuccoli, and S. Marinai, "Vehicle classification based on convolutional networks applied to fm-cw radar signals," *arXiv preprint arXiv:1710.05718*, 2017. [Online]. Available: https://arxiv.org/abs/1710.05718

[8] J. Liu *et al.*, "Background classification method based on deep learning for intelligent automotive radar target detection," *Future Generation Computer Systems*, vol. 94, pp. 524–535, 2019. [Online]. Available: https://doi.org/10.1016/j.future.2018.11.036

[9] S. Feintuch, H. H. Permuter, I. Bilik, and J. Tabrikian, "Neural network-based multi-target detection within correlated heavy-tailed clutter," *arXiv preprint arXiv:2210.12042*, 2022. [Online]. Available: https://arxiv.org/abs/2210.12042

[10] Defesa em Foco, "Embraer e exército apresentam o radar saber m200 vigilante," 2021, accessed: 2026-05-07. [Online]. Available: https://www.defesaemfoco.com.br/embraer-e-exercito-apresentam-o-radar-saber-m200-vigilante/

[11] M. A. Richards, *Fundamentals of Radar Signal Processing*, 2nd ed. New York, NY, USA: McGraw-Hill, 2014.

[12] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal, "Deep complex networks," *arXiv preprint arXiv:1705.09792*, 2018. [Online]. Available: https://arxiv.org/abs/1705.09792