

Paper Type: Original Article

RaDOD-Net: A Data-Contracted CNN–LSTM–Attention Architecture for Radar Enhancement and Airport Obstacle Detection

Mahmoud Alimoradi^{1*} , Setayesh Zahiri², Maryam Khani²

¹ Department of Computer Engineering, Islamic Azad University, Lahijan Branches, Lahijan, Iran; mahmoud.alimorady@gmail.com.

² Morvarid Intelligent Industrial Systems Research Group, Iran.

Citation:

Received: 03 January 2025

Revised: 16 March 2025

Accepted: 29 July 2025

Alimoradi, M., Zahiri, S., & Khani, M. (2026). RaDOD-Net: A data-contracted CNN–LSTM–attention architecture for radar enhancement and airport obstacle detection. *Computational Engineering and Technology Innovations*, 1(1), 30–40.

Abstract

Radar learning studies often place denoising, temporal feature extraction, attention, and target detection inside one pipeline while leaving the underlying radar product insufficiently specified. This omission is consequential: A polar intensity scan, a range–Doppler map, a range–azimuth map, and a raw complex radar cube do not support the same transformations, supervision, or claims. This paper reconstructs RaDOD-Net as a testable, product-aware architecture rather than as a report of unverified performance. The proposed system combines a shared convolutional encoder, patchwise causal LSTM states, local multi-head attention, a residual reconstruction head, and a calibrated detection head. Its central contribution is a radar data contract that binds every model component and metric to explicit axes, units, calibration, labels, temporal alignment, and provenance. A staged validation protocol separates signal restoration, target preservation, detection operating characteristics, domain transfer, and operational readiness. Oxford Radar RobotCar is treated as an urban polar-radar source that can support representation and transfer experiments after new annotation, but not as raw range–Doppler evidence or an airport benchmark. No numerical superiority is claimed. Instead, the paper provides tensor-level interfaces, leakage-resistant splits, matched baselines, uncertainty reporting, ablations, resource accounting, and evidence gates needed to test the architecture reproducibly. The result is a falsifiable research design for determining whether joint enhancement and detection improves weak-target surveillance without erasing target energy or inflating false alarms.

Keywords: Radar denoising, Obstacle detection, Range–Doppler, Polar radar, Causal temporal learning, Attention, Airport surveillance, Reproducibility.

1 | Introduction

Radar remains attractive for surveillance when illumination and visibility limit optical sensing, but the learning problem is not simply to make a radar image visually cleaner. Weak returns, interference, multipath, clutter, sidelobes, platform motion, and calibration drift can occupy overlapping regions of the measurement space. An enhancement network may suppress nuisance energy while also attenuating the small target that a detector

 Corresponding Author: mahmoud.alimorady@gmail.com

 <https://doi.org/10.48314/ceti.v3i1.82>



Licensee System Analytics. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0>).

must preserve. For airport monitoring, the cost of that error depends on target type, range, dwell time, protected zone, and the decision process that follows the alarm. Recent work has therefore examined radar support for bird, drone, and wake-vortex hazards, emphasizing continuous operation, rapid response, coverage, and automatic target recognition [1].

Deep Learning (DL) has expanded across radar waveform processing, interference mitigation, automatic target recognition, detection, and multimodal perception [2–5]. Convolutional autoencoders have been applied to range–Doppler denoising and FMCW interference mitigation [6–8]. Temporal models and attention have been used to exploit sequential radar structure [9], while deep detectors have been developed for pulse–Doppler UAV sensing [10], range–azimuth heatmaps [11], [12], and annotated range–angle–Doppler tensors [13], [14]. These studies demonstrate a broad design space, but they also show why the input product must be stated before the network: The axes and calibration of a radar tensor determine which physical effects are represented and which labels can be aligned with it.

Public radar datasets have accelerated repeatable comparison, yet their scientific roles differ. Oxford Radar RobotCar provides large-scale urban scanning-radar sequences and odometry [15]; CRUW supports radar object detection under camera-derived supervision [11], [12], [14]; CARRADA supplies range–angle–Doppler annotations [13]; RadarScenes emphasizes automotive radar point clouds [16]; nuScenes supplies multimodal autonomous-driving scenes [17]; and RADIATE focuses on adverse-weather radar perception [18]. None is, by itself, an airport obstacle benchmark. Treating these products as interchangeable creates an invalid chain from data to representation to claim.

RaDOD-Net addresses that chain through a data-contract-first architecture. The paper makes four bounded contributions : 1) it defines admissible paths for raw complex radar data and released polar power images, 2) it specifies a causal CNN–LSTM–attention network with dimension-matched reconstruction and detection heads, 3) it separates reconstruction quality from target-preservation and detection evidence, and 4) it introduces claim gates that prevent an urban representation result from being narrated as airport readiness. The work is intentionally methodological: it proposes what must be implemented and tested; it does not report trained weights, field trials, or numerical gains.

2 | Literature Architecture and the Unresolved Contract

2.1 | Radar Enhancement Is Not One Problem

The term denoising covers physically different tasks. Additive background noise, mutual FMCW interference, clutter, missing samples, and sensor artifacts require different corruption models and different evidence. Autoencoder-based studies have reconstructed range–Doppler products [6–8], fully convolutional and prior-guided models have treated interference mitigation [19], [20], and contrastive learning has been used to improve learned interference suppression [21]. These methods cannot be compared fairly when one receives raw complex samples and another receives a log-scaled map. A gain in image PSNR is not automatically a gain in detection probability, and phase preservation may matter even when magnitude images appear similar.

2.2 | Detection Depends on Representation and Supervision

Radar-only detection has evolved from classical thresholding and CFAR toward learned representations, cross-modal supervision, and dense tensor prediction. RODNet operated on range–azimuth heatmaps with camera-radar supervision [11], [12]; CARRADA paired camera information with radar tensor annotations [13]; and RADDet formulated detection on range–azimuth–Doppler data [22]. Dataset reviews have emphasized that representation, annotation granularity, sensor configuration, and scene diversity remain major sources of non-comparability [23]. Classical CA-CFAR remains important as an interpretable operating-point baseline, especially because learned systems must demonstrate their false-alarm behavior rather than only a summary accuracy [24].

2.3 | Temporal Learning and Attention Add Obligations

LSTM networks were designed to retain information over extended sequences [25], and attention allows a model to reweight relations among tokens or patches [26]. In radar, these mechanisms can aggregate weak evidence across time, but they introduce failure modes: motion misregistration can smear targets; bidirectional processing can violate online causality; global attention can be computationally prohibitive at native radar resolution; and a high-capacity temporal model can memorize session-specific clutter. Attention ConvLSTM results in traffic radar [9] motivate the combination, but an airport claim requires explicit temporal windows, registration, patch geometry, causal latency, and out-of-distribution tests.

2.4 | Airport Surveillance Changes the Endpoint

Bird–drone discrimination studies have used micro-Doppler signatures, CNNs, SNR-controlled evaluation, motion features, transfer learning, and radar-only pretraining [27–33]. Their endpoint is not merely a denoised frame but a decision under rare-target and low-false-alarm conditions. Airport surveillance additionally requires protected-zone geometry, target-specific detection range, alarm persistence, track confirmation, and safety integration. Consequently, automotive datasets can support pretraining or controlled representation studies, but airport transfer must be tested on relevant sensors, targets, clutter, weather, and operating procedures.

Table 1 organizes the literature by the scientific object actually evaluated. The unresolved gap is not the absence of another CNN–LSTM diagram. It is the absence of a single protocol that binds radar product, temporal architecture, enhancement target, detection threshold, and domain claim.

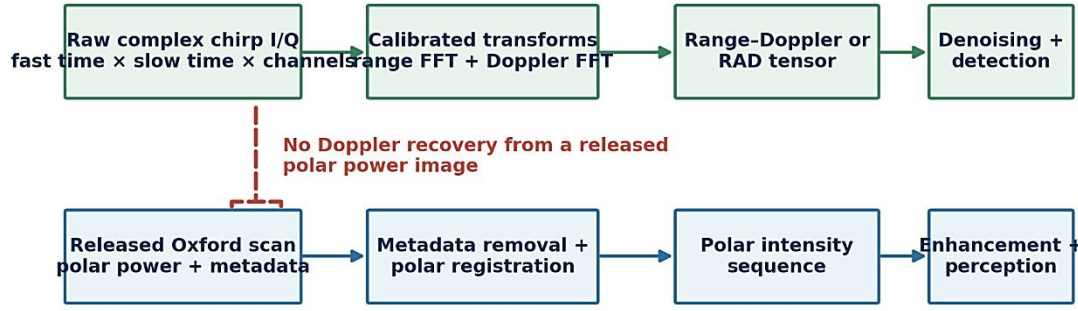
Table 1. Literature strands and the evidence that remains non-transferable.

Strand	Representative Studies	Established Object	What Does not Transfer Automatically
Radar DL surveys	[2–5]	Taxonomies of learning tasks, inputs, and architectures.	A survey-level opportunity does not validate a specific sensor, model, or deployment.
Enhancement / Interference mitigation	[6–8], [18–21]	Reconstruction or suppression under defined raw/map inputs.	Image fidelity does not establish target preservation or calibrated detection.
Radar object detection	[11–13], [16–18], [22], [23]	Detection on RA, RD/RAD, point-cloud, or multimodal products.	Metrics are not comparable across axes, labels, sensors, and matching rules.
Bird–drone surveillance	[1], [10], [27–31]	Classification or detection of small aerial targets.	Classification accuracy does not establish airport-wide false-alarm or response performance.
Sequence and attention models	[9], [25], [26]	Temporal memory and selective feature interaction.	Offline context and global attention do not imply causal, real-time feasibility.

3 | The Radar Data Contract

A RaDOD-Net experiment begins with a versioned data contract $C = (P, A, U, K, T, L, S)$, where P identifies the radar product; A lists tensor axes and ordering; U records physical units and scaling; K specifies calibration and preprocessing; T defines timestamps and temporal alignment; L defines labels and their mapping to radar support; and S records acquisition-session provenance. A result is admissible only for the contract under which it was produced. Changing from polar power scans to complex FMCW samples changes C and therefore defines a new experiment, not a preprocessing variant.

The radar product determines the admissible scientific question



A model may be shared across branches only after the input axes, units, labels, and loss are redefined.

Fig. 1. Radar-product contract.

Raw complex FMCW samples can support calibrated range–Doppler or range–angle–Doppler processing, whereas the released Oxford scan is a polar intensity product with embedded metadata. The two paths may inform a common architecture, but they cannot be conflated.

Let $X_1, \dots, X_t$ denote a causal sequence of maps with dimensions $H \times W \times C$. For raw data, these axes must be derived through documented windowing, transforms, calibration, and channel combination. For Oxford scans, the published polar PNG format and its metadata columns must be parsed according to the dataset documentation [15]. Normalization parameters μ and σ are estimated only on the training split and then frozen: $\tilde{X}_t = (X_t - \mu)/(\sigma + \epsilon)$. Any simulated corruption is applied after session splitting, with a saved seed and parameters. This ordering prevents information about held-out sessions from entering either normalization or corruption design.

Table 2. Admissible data paths for testing RaDOD-Net.

Input Product	Permitted Representation	Valid Immediate Task	Additional Evidence Required
Raw complex chirp I/Q	Calibrated RD or RAD tensor after fast-time, slow-time, and when supported angle processing.	Interference mitigation, reconstruction, detection.	Named source, chirp timing, calibration, windowing, target association, clean/reference protocol.
Oxford released scan	Polar power image after removal of embedded metadata and documented registration.	Enhancement, representation learning, odometry-adjacent perception after annotation.	New obstacle labels, target-support mapping, defensible corruption/reference, session-wise splits.
Automotive benchmark tensor	Dataset-specific RA, RD/RAD, point-cloud, or multimodal representation.	Pretraining, ablation, or benchmark detection.	Exact sensor and release version, coordinate conversion, annotation and metric definitions.
Airport radar acquisition	Product documented by the deployed sensor and collection protocol.	Domain validation and operational analysis.	Protected zones, target taxonomy, weather, surveillance hours, alarm and track definitions, safety review.

4 | RaDOD-Net Architecture

4.1 | Spatial Encoder and Causal Temporal State

For each normalized frame, a shared encoder E_0 produces $Z_t = E_0(\tilde{X}_t) \in \mathbb{R}^{h \times w \times d}$. Writing $M = hw$, feature patch m is $z_{t,m} \in \mathbb{R}^d$. A patchwise LSTM updates $(h_{t,m}, c_{t,m}) = \text{LSTM}_0(z_{t,m}, h_{t-1,m}, c_{t-1,m})$. The same patch index must correspond to the same physical support across frames; otherwise, the update requires registration or a

motion-conditioned correspondence. Hidden states are reset at session boundaries and never carried across train, validation, or test sequences.

4.2 | Local Attention Rather than Unconstrained Global Attention

Stack current hidden states as $H_t \in \mathbb{R}^{M \times d_h}$. For head i , $Q_i = H_t W_i Q$, $K_i = H_t W_i K$, $V_i = H_t W_i V$, and $A_i = \text{softmax}(Q_i K_i^T / \sqrt{d_k}) V_i$. Direct global attention costs $O(M^2)$ memory and is infeasible for large native scans. The executable design therefore uses windowed attention, shifted windows, axial attention, or a documented token reduction. The choice is part of the model specification because it changes receptive field, complexity, and boundary behavior.

4.3 | Dual Heads and Non-Destructive Enhancement

The fused feature F_t is decoded into a correction with the exact input shape, producing a residual output $\hat{S}_t = \tilde{X}_t + D\theta(F_t)$. A detection head $G\phi$ maps F_t and optionally $\hat{S}_t$ to per-cell or per-proposal scores p_t . The residual form does not guarantee target preservation, but it makes the identity mapping accessible and supports explicit penalties on target regions. When scores are interpreted probabilistically, calibration is estimated on validation data and evaluated on held-out sessions; modern networks can otherwise be overconfident even when classification accuracy is high [34].

RaDOD-Net as a causal, dimension-checked dual-task pipeline

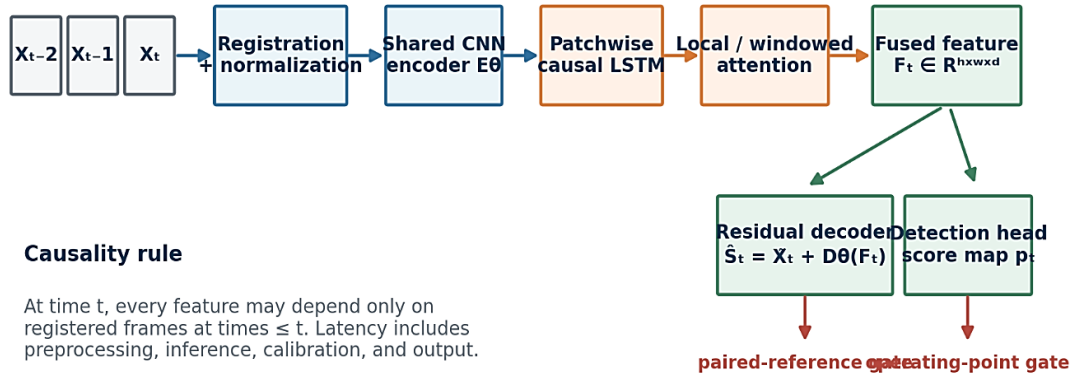


Fig. 2. Causal RaDOD-Net interface.

Registered radar sequences pass through a shared encoder, patchwise temporal state, and computationally bounded attention before branching into reconstruction and detection. Each head has a different evidence gate.

A joint loss may be written $L = \lambda_r L_r^{\text{ec}} + \lambda^d L^{\text{det}} + \lambda_p L_{\text{pr}}^{\text{es}} + \lambda_i L_{\text{it}}^{\text{emp}}$. Here L_r^{ec} is used only when a defensible clean or reference target exists; L^{det} is a class-imbalance-aware detection loss; $L_{\text{pr}}^{\text{es}}$ penalizes attenuation or displacement inside annotated target support; and $L_{\text{it}}^{\text{emp}}$ regularizes inconsistency across registered adjacent frames. Loss weights are selected without examining the test set. If a clean target cannot be justified, L_r^{ec} and reference-based PSNR are removed, and the task is described as enhancement with target-preservation and detection endpoints.

Table 3. Tensor and implementation ledger.

Object	Shape	Required Declaration	Principal Audit
Input X_t	$H \times W \times C$	Axis meaning, units, scale, product version.	No axis invented by preprocessing.
Encoded Z_t	$h \times w \times d$	Stride, padding, receptive field, parameter count.	Small-target support is not collapsed.
Temporal state H_t	$M \times d_h$	Window length, reset rule, registration, causality.	No future frame or cross-session state.
Attention A_t	$M \times d_h$	Locality pattern, heads, token count, memory.	Complexity measured at native input size.
Reconstruction $\hat{S}_t$	$H \times W \times C$	Output activation, residual scale, inverse normalization.	Target energy and morphology audited.
Detection p_t	grid or proposals	Coordinate system, matching, NMS, calibration.	Threshold fixed on validation only.

5 | Evaluation Architecture

5.1 | Staged Datasets and Leakage Control

Evaluation proceeds in stages rather than mixing all evidence into one headline table. Stage A verifies tensor correctness and reconstruction on a source with a justified target. Stage B evaluates detection on a dataset whose radar representation and annotations match the head. Stage C tests cross-dataset transfer without tuning on the destination test set. Stage D uses airport-relevant acquisitions for bird, drone, vehicle, and other obstacle classes. Oxford may contribute to Stage A or representation pretraining after new annotation, but its 32 urban traversals and polar scans do not satisfy Stage D [15].

Splits are made by acquisition session, traversal, day, or site before frame sampling. Adjacent frames, overlapping windows, repeated passes, and synthetic variants of the same clean frame remain in one partition. A manifest records file hashes, timestamps, sensor configuration, exclusions, label versions, and the exact split IDs. The split is frozen before architecture tuning. These controls matter because temporal radar sequences can yield deceptively high performance when neighboring frames leak across partitions.

5.2 | Baselines and Ablations

Every comparison receives the same product, split, label definition, threshold-selection rule, and preprocessing budget. Minimum baselines are: the detector on unenhanced input; a documented classical filtering or CFAR chain [24]; a convolutional autoencoder family [6–8]; a spatial-only CNN; a causal CNN–LSTM without attention; and the full model. Where the representation permits it, RODNet- or RADDet-family implementations provide detection context [11], [12], [22]. Ablations remove the residual connection, temporal state, attention, preservation term, and calibration separately. Parameter count, peak memory, multiply-accumulate estimate, and measured latency are reported for every learned baseline.

5.3 | Etrics Must Follow the Claim

Reconstruction metrics require matched references and a stated signal scale. PSNR is $10 \log_{10}(P^2/MSE)$, so P and the preprocessing scale must be reported. Signal gain in decibels is a paired difference under the same scene and region definition, not the nominal input SNR. Detection is reported as probability of detection or recall at prespecified false-alarm rates, plus precision–recall and receiver-operating curves. For surveillance, false alarms per scan, hour, or monitored area are more interpretable than an undefined percentage. Calibration is assessed with reliability diagrams and proper scores, not accuracy alone [34].

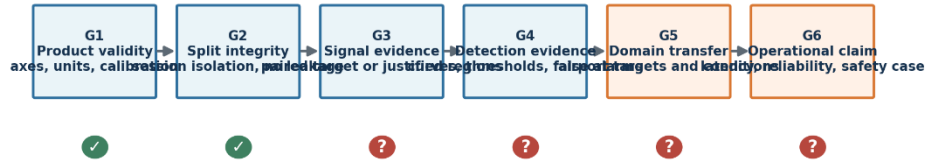
Table 4. Measures, operating rules, and archived evidence.

Claim	Primary Measure	Prespecified Rule	Evidence to Archive
Reference reconstruction	MAE/MSE; PSNR only with stated P.	Matched pairs and masks fixed before evaluation.	Pairs, scale, masks, per-session scores.
Target preservation	Energy ratio, peak displacement, support overlap.	Target regions defined from labels, not model output.	Before/After maps and region statistics.
Detection	Pd/Recall at fixed false-alarm points; PR/ROC curves.	Threshold chosen on validation data.	Scores, labels, matches, thresholds, curves.
False alarms	FP per Scan/Hour/Area and class of nuisance.	Denominator and surveillance duration declared.	Every negative opportunity and timestamp.
Calibration	Brier score, NLL, reliability curve.	Calibrator fitted on validation data only.	Pre/Post-calibration probabilities.
Deployment cost	End-to-end causal latency, throughput, memory.	Fixed hardware, map size, batch size, warm-up.	Profiling traces and deadline misses.

5.4 | Statistical Analysis and Robustness

Repeated training runs retain seeds and checkpoints. Confidence intervals use acquisition sessions, not individual frames, as the resampling unit when within-session dependence is substantial. Model comparisons report paired session-level differences with uncertainty and effect size; multiple pairwise comparisons are controlled rather than interpreted from isolated p-values [35–37]. Results are stratified by range, target size, weather, clutter, interference type, and sensor configuration. Robustness tests include missing frames, registration error, gain drift, unseen sites, changed radar hardware, and corruption strengths outside the training range.

Claims advance only when the matching evidence gate is passed



The present paper specifies G1-G2 and the tests for G3-G6; it does not report that the later gates have been passed.

Fig. 3. Evidence-gated claim hierarchy.

Product validity and split integrity are design prerequisites. Signal, detection, domain-transfer, and operational claims require progressively stronger evidence and cannot be inferred from an earlier gate.

6 | Airport Transfer, Safety, and Security Boundaries

Airport transfer begins with a hazard and operating concept, not with a generic dataset. The study must define protected zones, minimum detectable target classes, surveillance horizon, revisit time, permissible alarm delay, track-confirmation logic, handoff to other sensors, and the consequence of missed and false alarms. Bird–drone work shows that micro-Doppler, motion, and SNR affect discrimination [27–31], but the operational denominator is hours and area of surveillance under representative traffic and weather. Field evaluation should therefore report continuous negative exposure as carefully as positive target encounters.

The system boundary includes the radar, time synchronization, preprocessing, learned model, thresholding, tracker, alert interface, logging, and fallback. A model that misses a causal deadline or receives stale calibration

has failed even if its offline tensor metric is strong. Cybersecurity also matters: manipulated calibration, replayed radar frames, corrupted model files, or altered thresholds can redirect the decision chain. Versioned models, authenticated configuration, audit logs, health monitoring, and a fail-safe non-learning baseline should be included in deployment studies.

No safety certification or readiness claim follows from this architecture. The appropriate endpoint of the present work is a reproducible experiment that can falsify the proposed advantages. Operational adoption would require independent testing, change control, human-factors analysis, regulatory review, and site-specific safety assurance beyond the scope of a machine-learning benchmark.

7 | Limitations and Research Program

RaDOD-Net remains an untrained design specification. The paper does not supply weights, raw runs, airport data, or verified performance. The patchwise LSTM assumes meaningful correspondence after registration; strong target motion may require learned warping, tracking-conditioned state, or token association. Local attention controls complexity but can miss long-range interactions. Joint loss weighting may cause negative transfer between reconstruction and detection. A visually smoother output may be worse for detection, which is why target-preservation and operating-point tests are mandatory.

A disciplined research program would proceed through four studies. The first is a product-verification and simulation study with controlled corruption, matched pairs, and exact tensor tests. The second is a benchmark study using a representation-compatible public dataset, frozen splits, baselines, ablations, and repeated runs. The third is a cross-sensor transfer study that measures degradation without destination-test tuning. The fourth is a prospective airport campaign with synchronized radar, weather, ground truth, negative surveillance hours, causal profiling, and predeclared alarm rules. Evidence from each stage should be retained rather than replaced, so later claims inherit a traceable chain back to the input product.

8 | Conclusion

RaDOD-Net is recast as a data-contracted and evidence-gated radar learning architecture. Its CNN encoder, causal LSTM, bounded attention, residual reconstruction head, and calibrated detector form a coherent candidate pipeline only when the radar axes, units, calibration, temporal alignment, labels, and session provenance are explicit. This requirement prevents a released polar radar image from being treated as raw range-Doppler data and prevents a denoising metric from standing in for target detection. The proposed protocol connects each claim to its necessary experiment, from tensor validity and leakage control to false-alarm operating points, airport transfer, and causal resource measurement. The architecture is therefore presented as a falsifiable research program rather than as evidence of superiority. Its value will be established only if joint enhancement preserves weak targets and improves matched detection performance under representative radar products and airport conditions.

Authors' Contributions

All aspects of the research and manuscript preparation were carried out by the author. The author has read and approved the final version of the manuscript.

Funding

Not applicable.

Data Availability

All data supporting the reported findings in this research paper are provided within the manuscript.

Conflict of Interest

The author declares that they do not have any conflict of interest.

Consent for Publication

The author confirms consent for the publication of this work.

Ethics Approval and Consent to Participate

This article does not contain any studies with human participants performed by the author.

References

- [1] Gong, J., Yan, J., Kong, D., & Li, D. (2024). Enhancing airport airspace safety: A small radar system for mitigating bird, drone, and wake vortex hazards. *2024 AIAA DATC/IEEE 43rd digital avionics systems conference (DASC)* (pp. 1–10). IEEE. <https://doi.org/10.1109/DASC62030.2024.10749178>
- [2] Geng, Z., Yan, H., Zhang, J., & Zhu, D. (2021). Deep-learning for radar: A survey. *IEEE Access*, 9, 141800–141818. <https://doi.org/10.1109/ACCESS.2021.3119561>
- [3] Jiang, W., Ren, Y., Liu, Y., & Leng, J. (2022). Artificial neural networks and deep learning techniques applied to radar target detection: A review. *Electronics*, 11(1), 156. <https://doi.org/10.3390/electronics11010156>
- [4] Abdu, F. J., Zhang, Y., Fu, M., Li, Y., & Deng, Z. (2021). Application of deep learning on millimeter-wave radar signals: A review. *Sensors*, 21(6), 1951. <https://doi.org/10.3390/s21061951>
- [5] Jiang, W., Wang, Y., Li, Y., Lin, Y., & Shen, W. (2023). Radar target characterization and deep learning in radar automatic target recognition: A review. *Remote Sensing*, 15(15), 3742. <https://doi.org/10.3390/rs15153742>
- [6] de Oliveira, M. L. L., & Bekooij, M. J. G. (2020). Deep convolutional autoencoder applied for noise reduction in range-Doppler maps of FMCW radars. *IEEE radar conference 2020* (pp. 630–635). IEEE. <https://doi.org/10.1109/RADAR42522.2020.9114719>
- [7] Fuchs, J., Dubey, A., Lübke, M., Weigel, R., & Lurz, F. (2020). Automotive radar interference mitigation using a convolutional autoencoder. *2020 IEEE international radar conference (RADAR)* (pp. 315–320). IEEE. <https://doi.org/10.1109/RADAR42522.2020.9114641>
- [8] Chen, S., Taghia, J., Fei, T., Kühnau, U., Pohl, N., & Martin, R. (2021). A DNN autoencoder for automotive radar interference mitigation. *ICASSP 2021–2021 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 4065–4069). IEEE. <https://doi.org/10.1109/ICASSP39728.2021.9413619>
- [9] Jia, F., Tan, J., Lu, X., & Qian, J. (2023). Radar timing range-doppler spectral target detection based on attention ConvLSTM in traffic scenes. *Remote Sensing*, 15(17), 4150. <https://doi.org/10.3390/rs15174150>
- [10] Wang, C., Tian, J., Cao, J., & Wang, X. (2021). Deep learning-based UAV detection in pulse-Doppler radar. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–12. <https://doi.org/10.1109/TGRS.2021.3104907>
- [11] Wang, Y., Jiang, Z., Gao, X., Hwang, J. N., Xing, G., & Liu, H. (2021). Rodnet: Radar object detection using cross-modal supervision. *2021 IEEE winter conference on applications of computer vision (WACV)* (pp. 504–513). IEEE. <https://doi.org/10.1109/WACV48630.2021.00055>
- [12] Wang, Y., Jiang, Z., Li, Y., Hwang, J.-N., Xing, G., & Liu, H. (2021). RODNet: A real-time radar object detection network cross-supervised by camera-radar fused object 3D localization. *IEEE Journal of Selected Topics in Signal Processing*, 15(4), 954–967. <https://doi.org/10.1109/JSTSP.2021.3058895>
- [13] Ouaknine, A., Newson, A., Rebut, J., Tupin, F., & Pérez, P. (2021). Carrada dataset: Camera and automotive radar with range-angle-doppler annotations. *2020 25th international conference on pattern recognition (ICPR)* (pp. 5068–5075). IEEE. <https://doi.org/10.1109/ICPR48806.2021.9413181>

- [14] Wang, Y., Wang, G., Hsu, H. M., Liu, H., & Hwang, J. N. (2021). Rethinking of radar's role: A camera-radar dataset and systematic annotator via coordinate alignment. *2021 IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW)* (pp. 2809-2818). IEEE. <https://doi.org/10.1109/CVPRW53098.2021.00316>
- [15] Barnes, D., Gadd, M., Murcutt, P., Newman, P., & Posner, I. (2020). The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset. *2020 IEEE international conference on robotics and automation (ICRA)* (pp. 6433-6438). IEEE. <https://doi.org/10.1109/ICRA40945.2020.9196884>
- [16] Schumann, O., Hahn, M., Scheiner, N., Weishaupt, F., Tilly, J. F., Dickmann, J., & Wöhler, C. (2021). RadarScenes: A real-world radar point cloud data set for automotive applications. *2021 IEEE 24th international conference on information fusion (FUSION)* (pp. 1-8). IEEE. <https://doi.org/10.23919/FUSION49465.2021.9627037>
- [17] Caesar, H., Bankiti, V., Lang, A. H., Vora, S., Liong, V. E., Xu, Q., ... & Beijbom, O. (2020). Nuscenes: A multimodal dataset for autonomous driving. *2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR)* (pp. 11618-11628). IEEE. <https://doi.org/10.1109/CVPR42600.2020.01164>
- [18] Sheeny, M., De Pellegrin, E., Mukherjee, S., Ahrabian, A., Wang, S., & Wallace, A. (2021). Radiate: A radar dataset for automotive perception in bad weather. *2021 IEEE international conference on robotics and automation (ICRA)* (pp. 1-7). IEEE. <https://doi.org/10.1109/ICRA48506.2021.9562089>
- [19] Ristea, N. C., Anghel, A., & Ionescu, R. T. (2020). Fully convolutional neural networks for automotive radar interference mitigation. *2020 IEEE 92nd vehicular technology conference (VTC2020-Fall)* (pp. 1-5). IEEE. <https://doi.org/10.1109/VTC2020-Fall49728.2020.9348690>
- [20] Wang, J., Li, R., He, Y., & Yang, Y. (2022). Prior-guided deep interference mitigation for FMCW radars. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–16. <https://doi.org/10.1109/TGRS.2022.3211605>
- [21] Wang, J., Li, R., Zhang, X., & He, Y. (2023). Interference mitigation for automotive FMCW radar based on contrastive learning with dilated convolution. *IEEE Transactions on Intelligent Transportation Systems*, 25(1), 545–558. <https://doi.org/10.1109/TITS.2023.3306576>
- [22] Zhang, A., Nowruzi, F. E., & Laganieri, R. (2021). Raddet: Range-azimuth-doppler based radar object detection for dynamic road users. *2021 18th conference on robots and vision (CRV)* (pp. 95-102). IEEE. <https://doi.org/10.1109/CRV52889.2021.00021>
- [23] You, Y., Liu, W., Ze, Y., Li, Y. L., Wang, W., & Lu, C. (2022). Ukpagan: A general self-supervised keypoint detector. *2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR)* (pp. 17021-17030). IEEE. <https://doi.org/10.1109/CVPR52688.2022.01653>
- [24] Rohling, H. (1983). Radar CFAR thresholding in clutter and multiple target situations. *IEEE Transactions on Aerospace and Electronic Systems*, AES-19(4), 608–621. <https://doi.org/10.1109/TAES.1983.309350>
- [25] Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, 29(9), 2352–2449. https://doi.org/10.1162/neco_a_00990
- [26] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30. https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html
- [27] Dale, H., Baker, C., Antoniou, M., Jahangir, M., Atkinson, G., & Harman, S. (2022). SNR-dependent drone classification using convolutional neural networks. *IET Radar, Sonar & Navigation*, 16(1), 22–33. <https://doi.org/10.1049/rsn2.12161>
- [28] Rahman, S., & Robertson, D. A. (2020). Classification of drones and birds using convolutional neural networks applied to radar micro-Doppler spectrogram images. *IET Radar, Sonar & Navigation*, 14(5), 653–661. <https://doi.org/10.1049/iet-rsn.2019.0493>
- [29] Chen, X., Zhang, H., Song, J., Guan, J., Li, J., & He, Z. (2022). Micro-motion classification of flying bird and rotor drones via data augmentation and modified multi-scale CNN. *Remote Sensing*, 14(5), 1107. <https://doi.org/10.3390/rs14051107>

- [30] White, D., Jahangir, M., Baker, C., Antoniou, M., & Thiyagalingam, J. (2024). Comparing optical transfer learning and autoencoder pretraining for radar bird-drone discrimination. *2024 IEEE radar conference (RadarConf24)* (pp. 1-6). IEEE. <https://doi.org/10.1109/RadarConf2458775.2024.10548652>
- [31] Liu, J., Xu, Q. Y., & Chen, W. S. (2021). Classification of bird and drone targets based on motion characteristics and random forest model using surveillance radar data. *IEEE Access*, 9, 160135–160144. <https://doi.org/10.1109/ACCESS.2021.3130231>
- [32] Rahman, S., & Robertson, D. A. (2018). Radar micro-Doppler signatures of drones and birds at K-band and W-band. *Scientific Reports*, 8(1), 17396. <https://doi.org/10.1038/s41598-018-35880-9>
- [33] White, D. (2024). *Data-driven classification approaches for radar bird-drone surveillance* [Thesis]. <http://etheses.bham.ac.uk/id/eprint/15101>
- [34] Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. *International conference on machine learning* (pp. 1321-1330). PMLR. <https://proceedings.mlr.press/v70/guo17a.html>
- [35] Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7(Jan), 1–30. <https://www.jmlr.org/papers/volume7/demsar06a/demsar06a.pdf>
- [36] Garcia, S., & Herrera, F. (2008). An extension on "statistical comparisons of classifiers over multiple data sets" for all pairwise comparisons. *Journal of Machine Learning Research*, 9(12), 2677–2694. <https://www.jmlr.org/papers/volume9/garcia08a/garcia08a.pdf>
- [37] Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70. <https://www.jstor.org/stable/4615733>