Extended Abstract Track

Learnable Persistent Topology for Multi-View Radar Semantic Segmentation

Abstract

Objects in automotive radar produce sparse, fragmented echoes, making structural organisation central to interpreting Range–Doppler (RD), Range–Angle (RA), and Angle–Doppler (AD) views. Existing multi-view models capture which features group together, but not how connectivity and loop structure persist across scales. We introduce the first framework to learn cubical multifiltrations over a joint RD–RA–AD latent representation and integrate persistent H_0 and H_1 structure end-to-end into radar segmentation. Rather than relying on a fixed intensity-based construction, the filtration is learned, letting topology adapt to the fused representation; the resulting diagrams are vectorised, projected, and fused before RD and RA decoding. On CARRADA, the proposed method matches the strongest published methods in aggregate accuracy (63.85% RD, 44.70% RA mIoU) while achieving the highest car IoU in both views. Crucially, the gain is class-dependent: global persistence descriptors benefit spatially extended responses but not small, localised targets, a limitation of global conditioning rather than of persistence itself.

Keywords: radar semantic segmentation, persistent homology, topological deep learning, cubical persistence

1. Introduction

Automotive radar is robust under adverse conditions, but its sparse, noisy, and low-resolution measurements make semantic interpretation difficult (Yao et al., 2023; Kaul et al., 2020; Rebut et al., 2022). Multi-view methods combine complementary Range–Doppler (RD), Range–Angle (RA), and Angle–Doppler (AD) representations, with TMVA-Net (Ouaknine et al., 2021a) and TransRadar (Dalbah et al., 2024) demonstrating the value of cross-view modelling. HyperEcho (Khamis et al., 2026) subsequently showed that grouping fragmented echoes with hypergraphs captures higher-order reasoning. Building on this finding, HyperRadar (Zia et al., 2026b) learned higher-order structure within each view and aligned it across views using unbalanced optimal transport. These findings indicate that radar segmentation depends not only on feature content, but also on coherent structure within and across views, motivating topology as a complementary structural description.

Topological information has benefited deep learning and segmentation (Clough et al., 2020; Zia et al., 2023; Su et al., 2025). Whereas hypergraphs aggregate learned feature groups (Ali et al., 2026), persistent homology summarises global connectivity and loop structure across filtration scales (Zia et al., 2026a). This is attractive for radar because meaningful echoes are sparse and fragmented, yet their organisation can persist across filtration levels, while noise may form shorter-lived topological features. Cubical persistence is well suited to grid-structured radar representations (Rieck et al., 2020; Bleile et al., 2022), and learnable multiparameter vectorisations enable its integration into neural pipelines (Korkmaz et al., 2025).

© 2026 .

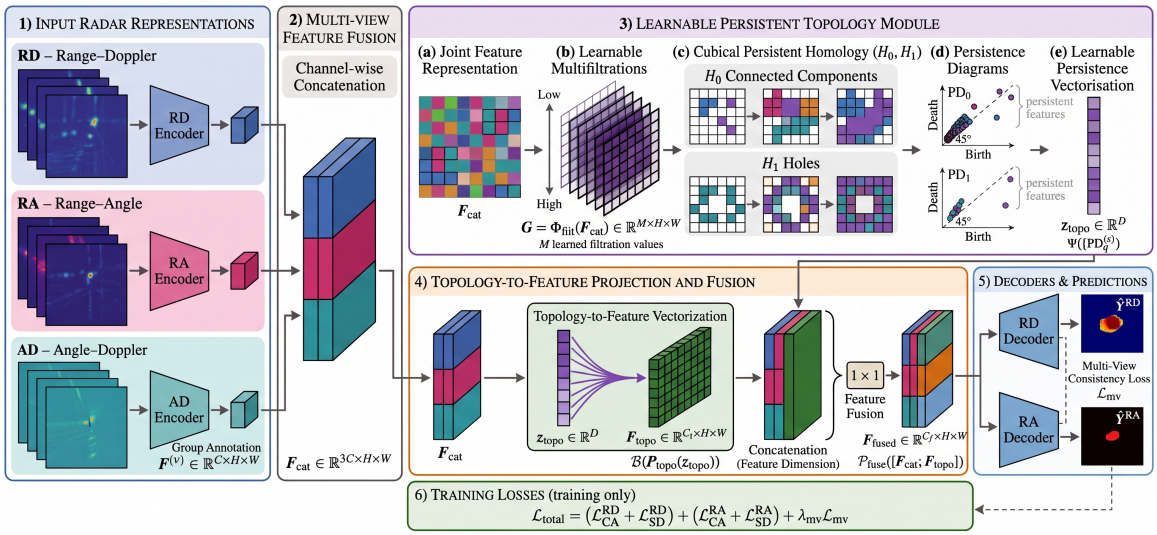


Figure 1: **Topology-guided multi-view radar segmentation.** RD, RA, and AD features are concatenated channel-wise, processed through learnable persistent topology, and fused with the original representation before RD and RA decoding.

We investigate persistent topology as a structural inductive bias for multi-view radar segmentation. Our framework extracts persistent H_0 and H_1 structures from the joint RD–RA–AD representation using learnable cubical multifiltrations and persistence vectorisation (Korkmaz et al., 2025). The descriptor is projected into feature space and fused with the original representation before RD and RA decoding. Experiments on CARRADA assess whether global topology complements conventional multi-view feature modelling.

2. Topology-Aware Multi-View Radar Segmentation

Multi-view encoding and fusion. We build on the TransRadar architecture (Dalbah et al., 2024), retaining its view-specific encoding, decoding, and optimisation design while introducing a persistent-homology pathway after multi-view fusion. Let $X^{(v)}$ denote the radar input for $v \in \text{RD}, \text{RA}, \text{AD}$. Dedicated encoders map each view to a common latent resolution, $F^{(v)} = E_v(X^{(v)}) \in \mathbb{R}^{C\times H\times W}$. The features are concatenated channel-wise as $F_{\text{cat}} = [F^{(\text{RD})}; F^{(\text{RA})}; F^{(\text{AD})}] \in \mathbb{R}^{3C\times H\times W}$, allowing persistent homology to characterise structure after the range, angle, and Doppler information has been combined.

Learnable persistent homology. The topology module takes F_{cat} as input and first learns a multiparameter cubical filtration $G = \Phi_{\text{filt}}(F_{\text{cat}}) \in \mathbb{R}^{M\times H\times W}$, where each spatial cell is assigned M learned filtration values. This allows the filtration to adapt to the joint radar representation rather than relying on a fixed intensity-based construction. Further background on cubical filtrations and persistent homology is provided in Appendix A. Following the differentiable multiparameter vectorisation strategy of (Korkmaz et al., 2025), the learned multifiltration is decomposed into single-parameter fil-

tration slices $G^{(s)}$. Cubical persistent homology is computed for each slice, retaining H_0 connected components and H_1 one-dimensional holes. Their birth–death pairs form $\mathrm{PD}_q^{(s)} = \{(b_{i,s}^q, d_{i,s}^q)\}_i$ for $q \in \{0,1\}$, where $d_{i,s}^q - b_{i,s}^q$ denotes the persistence of a topological feature. The resulting variable-size diagrams are aggregated by a learnable persistence vectorisation, $z_{\mathrm{topo}} = \Psi(\{\mathrm{PD}_q^{(s)}\}_{q,s}) \in \mathbb{R}^D$, yielding a fixed-dimensional descriptor of persistent connectivity and hole structure in the joint RD–RA–AD latent space.

Topology-to-feature projection and fusion. Because z_{topo} is a global descriptor whereas segmentation requires spatial features, it is projected to C_t channels and broadcast over the latent grid as $F_{\mathrm{topo}} = \mathcal{B}(P_{\mathrm{topo}}(z_{\mathrm{topo}})) \in \mathbb{R}^{C_t \times H \times W}$. The topology-derived features are concatenated with the original representation and combined using a 1×1 convolution, $F_{\mathrm{fused}} = P_{\mathrm{fuse}}([F_{\mathrm{cat}}; F_{\mathrm{topo}}]) \in \mathbb{R}^{C_f \times H \times W}$. Persistent homology therefore provides global structural conditioning, while F_{cat} retains spatially resolved radar information.

Semantic decoding and optimisation. Parallel decoders produce $\hat{Y}^{\mathrm{RD}} = D_{\mathrm{RD}}(F_{\mathrm{fused}})$ and $\hat{Y}^{\mathrm{RA}} = D_{\mathrm{RA}}(F_{\mathrm{fused}})$. The model is jointly optimised using class-agnostic object ($\mathcal{L}_{\mathrm{CA}}$) and soft-Dice ($\mathcal{L}_{\mathrm{SD}}$) losses for both views, together with the multi-view range-consistency loss $\mathcal{L}_{\mathrm{mv}}$: $\mathcal{L}_{\mathrm{total}} = (\mathcal{L}_{\mathrm{CA}}^{\mathrm{RD}} + \mathcal{L}_{\mathrm{SD}}^{\mathrm{RD}}) + (\mathcal{L}_{\mathrm{CA}}^{\mathrm{RA}} + \mathcal{L}_{\mathrm{SD}}^{\mathrm{RA}}) + \lambda_{\mathrm{mv}}\mathcal{L}_{\mathrm{mv}}$. Here, $\mathcal{L}_{\mathrm{mv}}$ encourages consistent class responses along the range dimension shared by the RD and RA predictions.

3. Experiments and Results

We evaluate on CARRADA (Ouaknine et al., 2021b), which provides synchronised RD, RA, and AD radar representations with semantic annotations. Following standard multi-view radar evaluation (Dalbah et al., 2024), we report mIoU and mDice for RD and RA segmentation, together with foreground class IoU for pedestrian, cyclist, and car regions. We compare against TMVA-Net (Ouaknine et al., 2021a), TransRadar (Dalbah et al., 2024), and recent structurally motivated approaches, including HyperEcho (Khamis et al., 2026), which models higher-order relationships using hypergraphs, and HyperRadar (Zia et al., 2026b), which learns structurally consistent multi-view representations.

Quantitative analysis. Table 1 reports results under consistent experimental settings. Our topology-aware model achieves the best mIoU in RD (63.85%) and RA (44.70%), the best RA mDice (54.65%), and the second-best RD mDice (74.70%). It also performs the best car IoU in RD (64.2%) and RA (39.1%), RD cyclist Dice (49.1%), and RA car Dice (55.42%). HyperRadar performs better on several remaining metrics, while HyperEcho leads RD pedestrian and RA cyclist segmentation. Overall, the aggregate differences from HyperRadar are marginal (+0.05 RD mIoU, +0.30 RA mIoU), while the clearest and most consistent gain is on the car class in both views. This pattern indicates a class-dependent rather than uniform effect.

Qualitative analysis. Figure 2 illustrates topology-aware modelling in two CARRADA scenes. In these examples, our predictions preserve more spatially complete car regions, while several competing outputs contain smaller, fragmented, or missed responses. The first scene also shows multiple foreground targets retained across RD and RA, whereas the second shows improved preservation of the dominant car region. These observations

Table 1: Semantic segmentation performance on CARRADA. **Bold** and **blue bold** denote the best and second-best results, respectively, within each view; ties are retained.

View	Method	IoU (%)					Dice (%)				
		Bkg.	Ped.	Cycl.	Car	mIoU	Bkg.	Ped.	Cycl.	Car	mDice
RD	TMVA-Net (Ouaknine et al., 2021a)	**99.7**	52.6	29.0	53.4	58.7	**99.8**	68.9	45.0	69.6	70.9
	TransRadar (Dalbah et al., 2024)	**99.9**	56.7	30.2	61.7	62.1	**99.8**	72.4	46.4	76.3	73.4
	HyperEcho (Khamis et al., 2026)	**99.7**	**60.33**	31.86	57.7	62.42	**99.8**	**75.26**	48.32	73.21	74.16
	HyperRadar (Zia et al., 2026b)	**99.7**	59.2	**32.4**	**63.8**	**63.8**	**99.8**	**74.5**	**48.9**	**77.6**	**75.2**
	Topology-Aware (Ours)	**99.7**	**59.3**	**32.2**	**64.2**	**63.85**	**99.8**	73.3	**49.1**	**76.6**	**74.70**
RA	TMVA-Net (Ouaknine et al., 2021a)	99.8	26.0	8.6	30.7	41.3	**99.9**	41.3	**15.9**	47.0	51.0
	TransRadar (Dalbah et al., 2024)	**99.9**	29.9	6.5	35.3	42.9	**99.9**	46.0	12.2	52.2	52.6
	HyperEcho (Khamis et al., 2026)	**99.86**	26.12	**11.96**	31.25	42.30	**99.93**	41.43	**23.23**	47.62	52.50
	HyperRadar (Zia et al., 2026b)	**99.9**	**31.8**	8.4	**37.5**	**44.4**	**99.9**	**48.2**	14.6	**54.4**	**54.3**
	Topology-Aware (Ours)	**99.9**	**30.4**	**9.4**	**39.1**	**44.70**	**99.9**	**48.07**	15.21	**55.42**	**54.65**

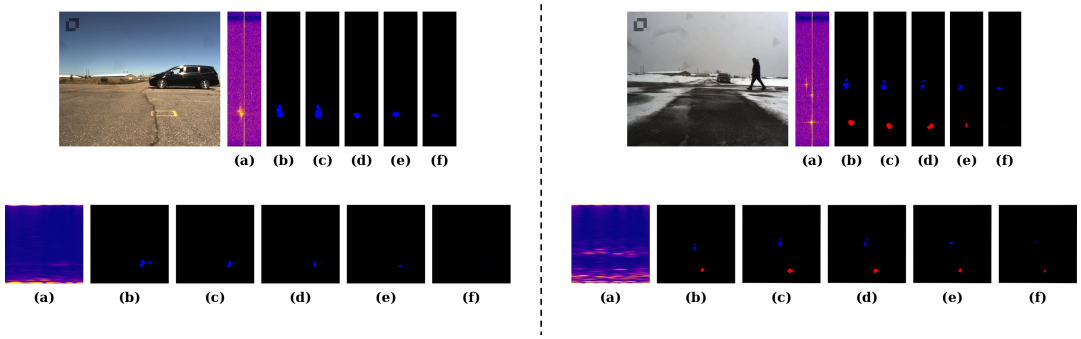


Figure 2: **Qualitative comparison on CARRADA.** Two test scenes are shown side by side, with RD (top) and RA (bottom) views: (a) radar input, (b) ground truth, (c) **Ours**, (d) TransRadar (Dalbah et al., 2024), (e) TMVA-Net (Ouaknine et al., 2021a), and (f) MVNet (Ouaknine et al., 2021a).

are consistent with the class-wise car results in Table 1, but are intended as qualitative examples rather than evidence of a uniform improvement.

4. Discussion and Conclusion

The results reveal a class-dependent rather than uniform effect. The topology-guided model achieves the strongest car IoU in both views, while aggregate performance is essentially unchanged relative to the strongest prior method (+0.05 RD and +0.30 RA mIoU). The contrast between a clear class-level gain and a flat aggregate suggests that global persistence descriptors may favour spatially extended responses over small, highly localised targets. The current broadcast design provides sample-level structural context without indicating where individual components or holes occur, which may explain this behaviour and limit fine-grained localisation. Persistent homology should therefore be viewed as a complementary global prior rather than a replacement for local and cross-view feature modelling. Future work will develop spatially adaptive topological features and radar-geometry-aware filtrations, while separately evaluating the contributions of H_0 and H_1.

References

Usman Ali, Ali Zia, Waqas Ali, Umer Ramzan, Abdul Rehman, Muhammad Tayyab Chaudhry, and Wei Xiang. Hypergraph contrastive sensor fusion for multimodal fault diagnosis in induction motors. *IEEE Sensors Journal*, 26(4):5829–5840, 2026. doi: 10.1109/JSEN.2025.3648413.

Bea Bleile, Adélie Garin, Teresa Heiss, Kelly Maggs, and Vanessa Robins. The persistent homology of dual digital image constructions. In *Research in Computational Topology 2*, pages 1–26. Springer, 2022.

James R Clough, Nicholas Byrne, Ilkay Oksuz, Veronika A Zimmer, Julia A Schnabel, and Andrew P King. A topological loss function for deep-learning based image segmentation using persistent homology. *IEEE transactions on pattern analysis and machine intelligence*, 44(12):8766–8778, 2020.

Yahia Dalbah, Jean Lahoud, and Hisham Cholakkal. Transradar: Adaptive-directional transformer for real-time multi-view radar semantic segmentation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 353–362, 2024.

Prannay Kaul, Daniele De Martini, Matthew Gadd, and Paul Newman. Rss-net: Weakly-supervised multi-class semantic segmentation with fmcw radar. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 431–436. IEEE, 2020.

Abdelwahed Khamis, Muhammad Umer Ramzan, Usman Ali, and Ali Zia. Hyperecho: Structured higher order representation for automotive radar semantic segmentation. In *2026 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, pages 1–6. IEEE, 2026.

Caner Korkmaz, Brighton Nuwagira, Baris Coskunuzer, and Tolga Birdal. Cumperlay: Learning cubical multiparameter persistence vectorizations. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 27084–27094, 2025.

Arthur Ouaknine, Alasdair Newson, Patrick Pérez, Florence Tupin, and Julien Rebut. Multi-view radar semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 15671–15680, 2021a.

Arthur Ouaknine, Alasdair Newson, Julien Rebut, Florence Tupin, and Patrick Pérez. Carrada dataset: Camera and automotive radar with range-angle-doppler annotations. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 5068–5075. IEEE, 2021b.

Julien Rebut, Arthur Ouaknine, Waqas Malik, and Patrick Pérez. Raw high-definition radar for multi-task learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17021–17030, 2022.

Bastian Rieck, Tristan Yates, Christian Bock, Karsten Borgwardt, Guy Wolf, Nicholas Turk-Browne, and Smita Krishnaswamy. Uncovering the topology of time-varying fmri

data using cubical persistence. *Advances in neural information processing systems*, 33: 6900–6912, 2020.

Zhe Su, Xiang Liu, Layal Bou Hamdan, Vasileios Maroulas, Jie Wu, Gunnar Carlsson, and Guo-Wei Wei. Topological data analysis and topological deep learning beyond persistent homology: a review. *Artificial Intelligence Review*, 2025.

Shanliang Yao, Runwei Guan, Xiaoyu Huang, Zhuoxiao Li, Xiangyu Sha, Yong Yue, Eng Gee Lim, Hyungjoon Seo, Ka Lok Man, Xiaohui Zhu, et al. Radar-camera fusion for object detection and semantic segmentation in autonomous driving: A comprehensive review. *IEEE Transactions on Intelligent Vehicles*, 9(1):2094–2128, 2023.

A Zia, A Khamis, J Nichols, Z Hayder, V Rolland, and L Peterssonet. Topological deep learning: A review of an emerging paradigm. arxiv. *arXiv preprint arXiv:2302.03836*, 2023.

Ali Zia, Usman Ali, Abdul Rehman, Umer Ramzan, Kang Han, Muhammad Faheem, Shahnawaz Qureshi, and Wei Xiang. Learning topology-aware representations via test-time adaptation for anomaly segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2026a. doi: 10.1109/TPAMI.2026.3710196. Early Access.

Ali Zia, Muhammad Umer Ramzan, Abdelwahed Khamis, Usman Ali, and Abdul Rehman. Learning structurally consistent representations for multi-view radar semantic segmentation. *arXiv preprint arXiv:2606.31609*, 2026b.

Appendix A. Topological Background

This appendix summarises the topological construction used in the system and provides additional detail for the learnable persistent-topology module described in Sec. 2.

Cubical multifiltration. Topology is extracted from the joint multi-view latent representation $F_{\text{cat}} \in \mathbb{R}^{3C\times H\times W}$ rather than from the individual radar views. Let $\Omega = \{1,\dots,H\}\times\{1,\dots,W\} \subset \mathbb{Z}^2$ denote its spatial grid. A learnable filtration mapping produces

$$G = \Phi_{\text{filt}}(F_{\text{cat}}) \in \mathbb{R}^{M\times H\times W}, \tag{1}$$

such that every spatial location $u \in \Omega$ is assigned an M-dimensional filtration value $G(u) = [G_1(u),\dots,G_M(u)]^\top \in \mathbb{R}^M$. Together with the regular grid structure, these values define a cubical multifiltration: cells are progressively included according to the learned filtration coordinates rather than a hand-crafted radar-intensity function.

Single-parameter filtration slices. Ordinary persistence diagrams are obtained by reducing the learned multiparameter filtration to a collection of S single-parameter slices, following the multiparameter persistence-vectorisation strategy of (Korkmaz et al., 2025). Let $\mathcal{S}_s$ denote the s-th learnable slicing operator. The corresponding scalar filtration map is

$$G^{(s)}(u) = \mathcal{S}_s(G(u)), \qquad s = 1,\dots,S. \tag{2}$$

For ordered filtration values $\tau_1 < \tau_2 < \cdots < \tau_N$, each slice induces a nested sequence of cubical subcomplexes

$$K_1^{(s)} \subseteq K_2^{(s)} \subseteq \cdots \subseteq K_N^{(s)}, \qquad K_t^{(s)} = \{u \in \Omega \mid G^{(s)}(u) \le \tau_t\}. \tag{3}$$

Hence, each $G^{(s)}$ provides a different one-dimensional view of the learned multifiltration.

Cubical persistent homology. Persistent homology tracks the appearance and disappearance of topological structures along each filtration in Eq. 3. We retain homology dimensions $q \in \{0,1\}$, where H_0 represents connected components and H_1 represents one-dimensional holes. For slice s, the corresponding persistence diagram is

$$\text{PD}_q^{(s)} = \left\{(b_{i,s}^q, d_{i,s}^q)\right\}_i, \qquad q \in \{0,1\}, \tag{4}$$

where $b_{i,s}^q$ and $d_{i,s}^q$ denote the birth and death values of the i-th topological feature. Its persistence is

$$\pi_{i,s}^q = d_{i,s}^q - b_{i,s}^q. \tag{5}$$

Features with larger persistence remain present over a wider filtration interval and therefore represent more stable topological structure.

Learnable persistence vectorisation. Because each $\text{PD}_q^{(s)}$ contains a variable number of birth–death pairs, the diagrams cannot be directly concatenated with neural feature tensors. A learnable persistence-vectorisation operator aggregates the H_0 and H_1 information across all filtration slices into a fixed-dimensional descriptor,

$$z_{\text{topo}} = \Psi\Big(\{\text{PD}_q^{(s)}\}_{q\in\{0,1\},\, s=1}^{S}\Big) \in \mathbb{R}^D. \tag{6}$$

Thus, z_{topo} summarises persistent connectivity and hole structure present in the joint RD–RA–AD latent representation while providing a fixed-size representation suitable for subsequent neural processing.

Topology-to-feature projection. The descriptor z_{topo} contains global structural information, whereas semantic decoding operates on spatial feature maps. It is therefore projected to C_t feature channels,

$$p_{\text{topo}} = P_{\text{topo}}(z_{\text{topo}}) \in \mathbb{R}^{C_t}, \tag{7}$$

and broadcast over the common latent grid,

$$F_{\text{topo}} = \mathcal{B}(p_{\text{topo}}) \in \mathbb{R}^{C_t \times H \times W}. \tag{8}$$

The topology-derived features are concatenated with the original multi-view representation and fused using a 1×1 convolution,

$$F_{\text{fused}} = P_{\text{fuse}}\left([F_{\text{cat}}; F_{\text{topo}}]\right) \in \mathbb{R}^{C_f \times H \times W}. \tag{9}$$

This design retains the spatially resolved information in F_{cat} while allowing the persistent topological descriptor to provide complementary global structural information before RD and RA semantic decoding.