Topology-Guided Test-Time Adaptation via Persistent Homology

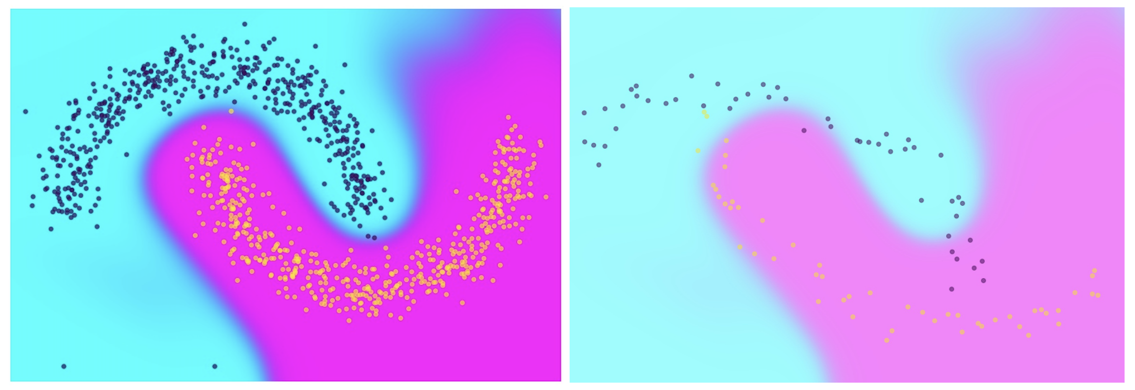
From Affective Behavior Analysis to Autonomous Driving

Onur Cezmi Mutlu* • Mohammadmahdi Honarmand* • Parnian Azizian • Saimourya Surabhi • Dennis P. Wall

Stanford University – {cezmi, mhonar, azizian, mourya, dpwall}@stanford.edu – 10th ABAW Workshop, CVPR 2026 – *Equal contribution

The Problem & Our Insight

Blind test-time adaptation can hurt performance. Existing TTA methods (TENT, self-training, pseudo-labeling) assume adaptation is *always* beneficial — a risky assumption for affective computing and autonomous driving.



Source: clean separation *Shifted: model fails*

Not *when* has data shifted? but *will* adapting help?

The **topology of intermediate activations** reveals whether TTA will succeed: *stable* multi-scale topology ⇒ adapt safely; *fragile* topology ⇒ skip. We propose ***Informed Adaptation*** — a topology-aware decision layer predicting adaptation success *before* any update; adapter-agnostic, $<1\%$ overhead.

Contributions

1. **First** topology-guided *decision* framework for TTA.
2. **Universal wrapper** compatible with any TTA method.
3. Validated across video emotion (AffWild2) & spatial detection (SHIFT); CNN + Transformer backbones.
4. **Zero observed degradations**; smaller models benefit disproportionately.

Algorithm 1: Topology-Guided TTA

Input: batch $\mathcal{X}$, model f, TTA method $\mathcal{A}$, classifier g.

1. Extract activations; construct graph G.

2. Compute persistence diagrams via filtration.

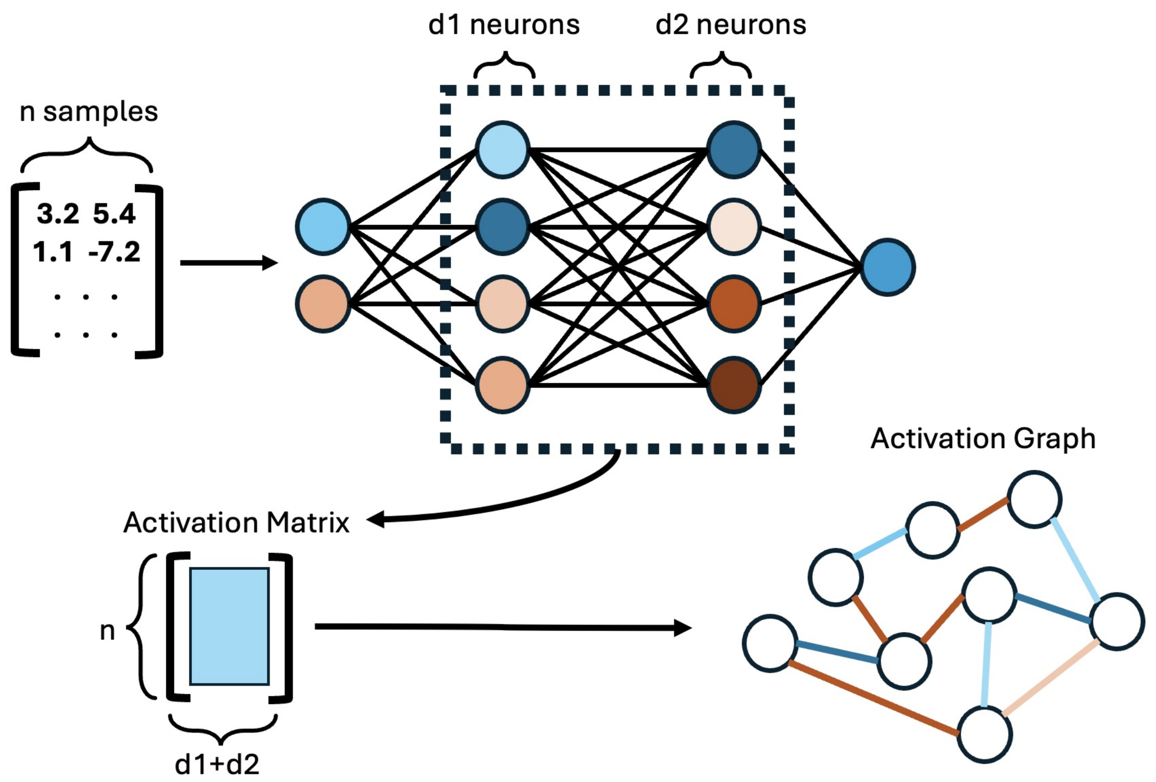
3. Extract 30-dim feature vector $\mathbf{z} \in \mathbb{R}^{30}$.

4. if $g(\mathbf{z}) = 1$ **return** $\mathcal{A}(f, \mathcal{X})(\mathcal{X})$ *(apply TTA)*

else return $f(\mathcal{X})$ *(skip adaptation)*

For video, decisions are made per 10-frame window.

Pipeline



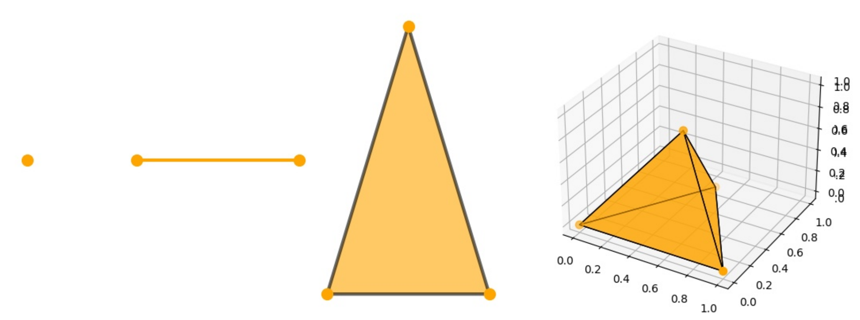
Activations → normalized feature vectors → weighted activation graph (cosine similarity) → persistent homology → 30-dim descriptor → classifier.

Normalize across the batch, then connect all pairs of nodes with cosine-similarity edge weights:

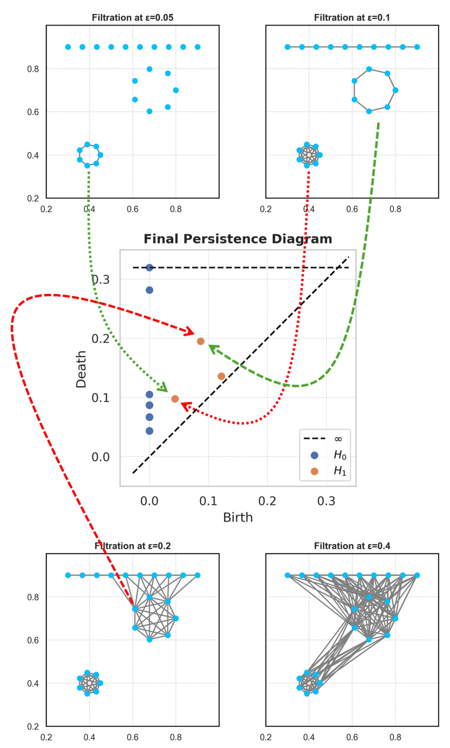
$$\widetilde{\mathbf{h}}_i = \frac{\mathbf{h}_i - \mu(\mathbf{H})}{\sigma(\mathbf{H})} \qquad w_{ij} = \frac{\widetilde{\mathbf{h}}_i \cdot \widetilde{\mathbf{h}}_j}{\|\widetilde{\mathbf{h}}_i\|\,\|\widetilde{\mathbf{h}}_j\|}$$

Graphs capped at 2500 vertices for tractable persistent homology.

Persistent Homology



0-, 1-, 2-, 3-simplices: building blocks of complexes.

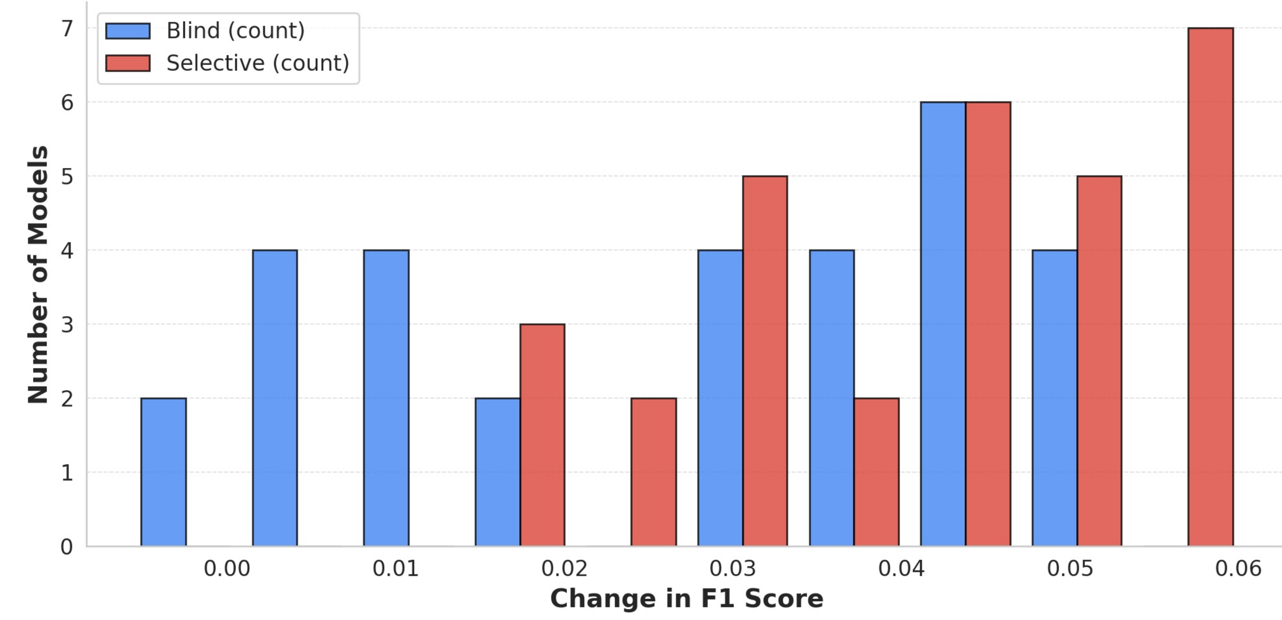


Vietoris–Rips filtration $\varnothing = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_m = K$ → *persistence diagram* of H_0 (clusters) & H_1 (loops); off-diagonal points mark stable structures.

4 attributes × 7 statistics + 2 Betti numbers = 30-dim descriptor $\mathbf{z} \in \mathbb{R}^{30}$ **→ XGBoost classifier** $g : \mathbb{R}^{30} \to \{\text{adapt}, \text{skip}\}$

AffWild2 — Emotion Recognition In-the-Wild

Setup: 558 emotion-recognition videos; backbones ResNet-18d, ResNet-50d, MobileNetV2; underlying TTA = TempT (temporal-consistency entropy minimization).

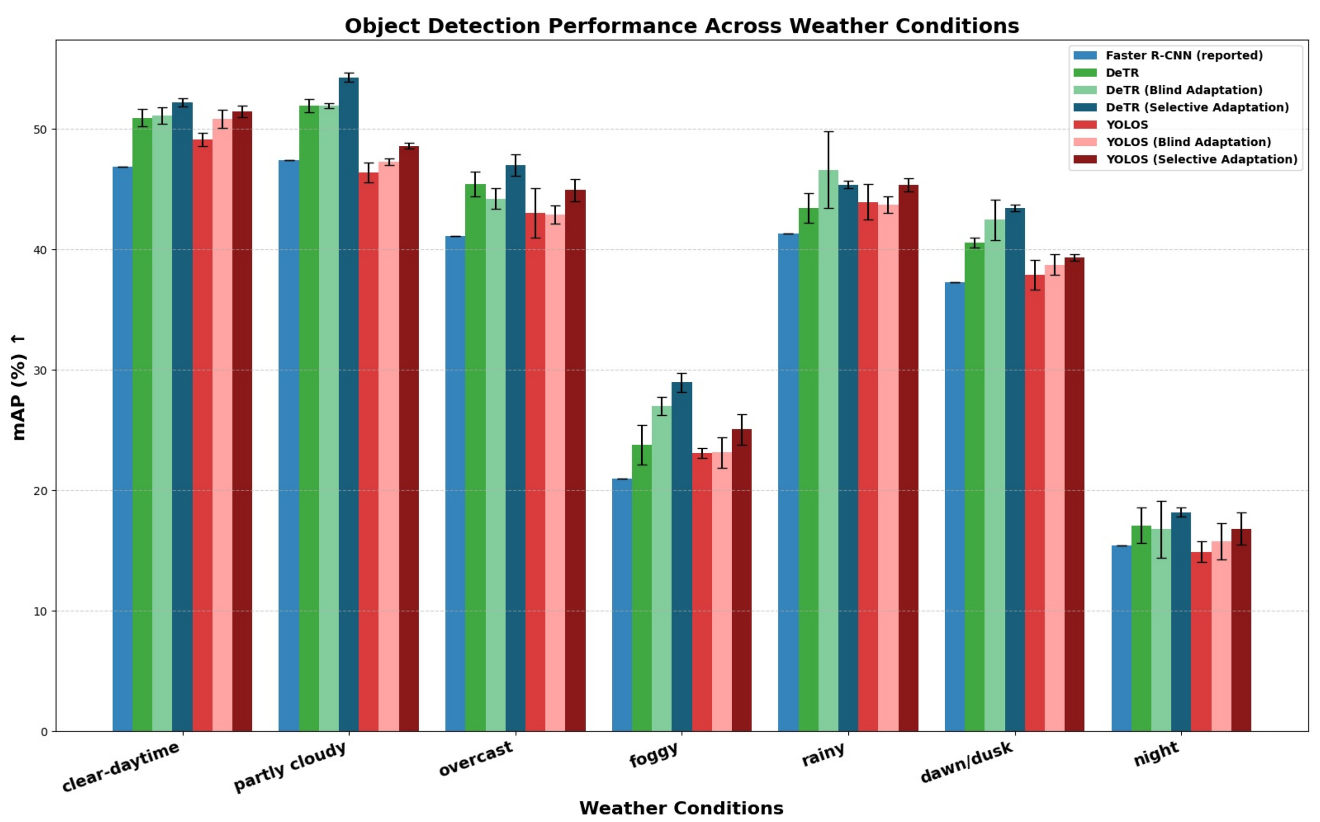


Distribution of per-video F1 changes: topology-guided selective adaptation (red) yields larger and more consistent positive gains than blind adaptation (blue).

- $+2.5\%$ avg. F1 improvement vs. blind adaptation
- **Zero** observed degradations
- Decision overhead $<1\%$ of TempT runtime

SHIFT — Object Detection (Autonomous Driving)

Setup: 4,250 sequences (2M+ frames); YOLOS & DeTR transformers vs. Faster R-CNN baseline; mAP@0.5 across 7 weather conditions.



Selective adaptation (dark bars) dominates baseline & blind variants in every weather.

- **Systematic mAP gains** across all 7 weathers
- Blind adaptation *inconsistent* (hurts in fog, rain)
- **Smaller models benefit most**: YOLOS+sel. ≈ DeTR+blind

Headline Results

+2.5%

avg F1 gain on AffWild2

0

observed degradations

<1%

compute overhead

Takeaways

- Multi-scale topology of activations is a **reliable, interpretable** signal for safe adaptation decisions.
- Same 30-dim descriptor transfers across domains and architecture families.
- Eliminates the negative tail of TTA failures — critical for safety-sensitive deployment.

Limitations & Future

- Binary adapt/skip → multi-class (*how much*, *which adapter*).
- H_0–H_2 via Vietoris–Rips only; richer complexes may help.
- Long-horizon continual streams; integrate with causal / uncertainty-aware criteria.

Acknowledgments

We thank all members of the Wall Lab. Supported by NIH (R01LM014342, R01LM013364, DP2EB035858), NSF Award 2500517, and the Stanford Medicine Center for Digital Health.

Key References

Wang et al. **TENT**, ICLR 2021.
Mutlu et al. **TempT**, CVPRW 2023.
Carlsson, **Topology and Data**, Bull. AMS 2009.
Edelsbrunner & Harer, **Pers. Hom. Survey**, 2008.
Corneanu et al. **Test Error w/o Test Set**, CVPR 2020.
Sun et al. **SHIFT**, CVPR 2022.
Kollias & Zafeiriou, **ABAW**.