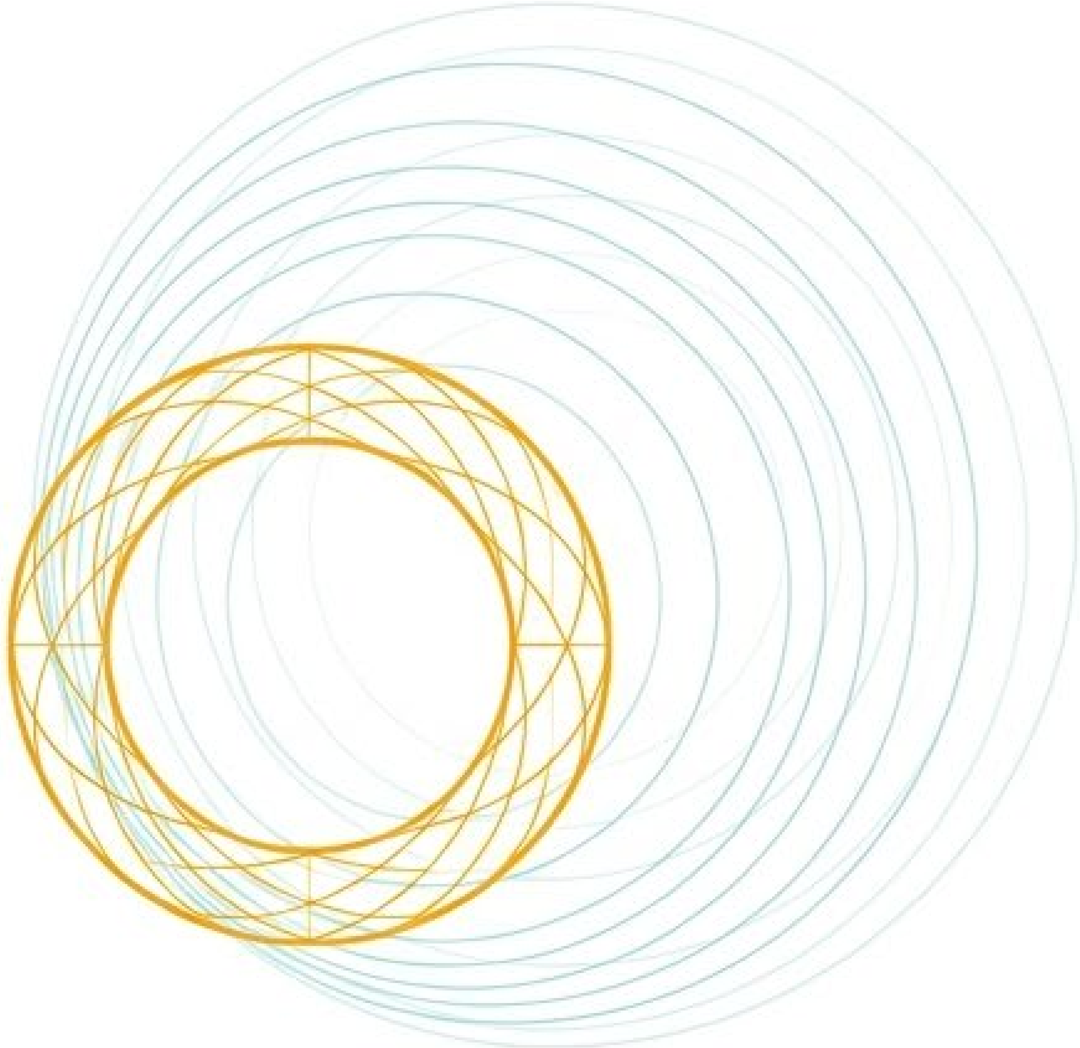
# BiSTS: Biphasic Spatial-Temporal Synergy for Approaching Motion Recognition

A bio-inspired, mathematically guaranteed architecture for robust edge-vision collision detection.



# The fundamental bottleneck in bio-inspired edge vision

LGMD architectures offer immense **computational efficiency** but **fail catastrophically in dynamic environments.**

## Biological Efficiency

- Learning-free structural design
- Exceptionally low computational footprint
- Hardware-friendly membrane potential tracking

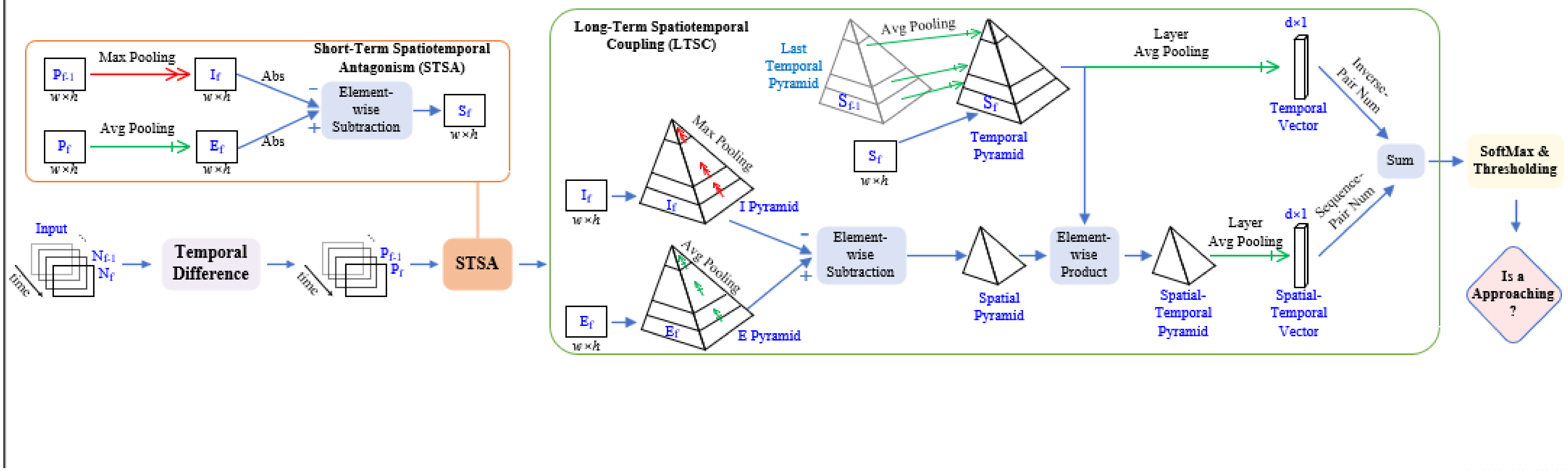
## Environmental Vulnerability

- Severe susceptibility to high-frequency visual noise
- Complete breakdown under low-contrast
- Inability to distinguish true looming objects from lateral translation or receding motions

Result: Unacceptable false alarm rates in real-world autonomous navigation.

# A biphasic architecture harmonizing antagonism and coupling

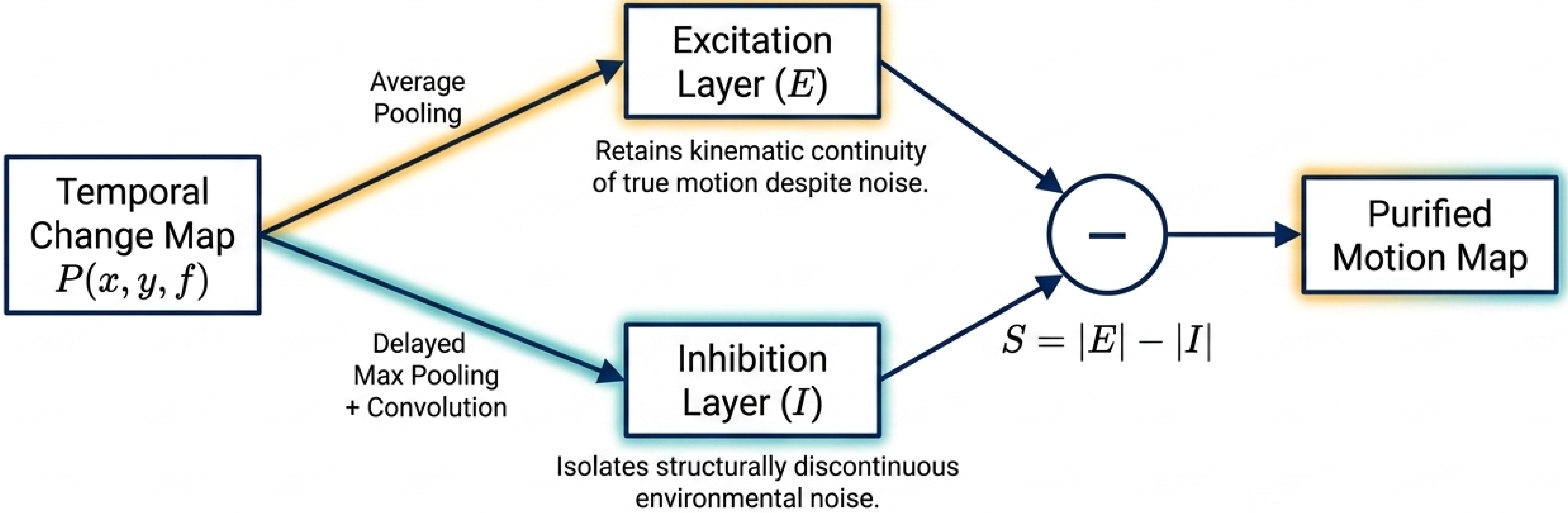
BiSTS explicitly decouples short-term noise suppression from long-term spatiotemporal pattern recognition.




NotebookLM

# Phase 1: Short-Term Spatiotemporal Antagonism (STSA)

Exploiting the complementary properties of average and max pooling to filter localized interference.




By correlating E and I spatially but antagonizing them temporally, STSA preserves the diffusion pattern of a looming object while nullifying random spatial spikes.

# A mathematical guarantee against visual noise

Unlike standard models, the STSA output expects a strictly negative baseline response to random interference.

## Gaussian Noise (Theorem B.3)

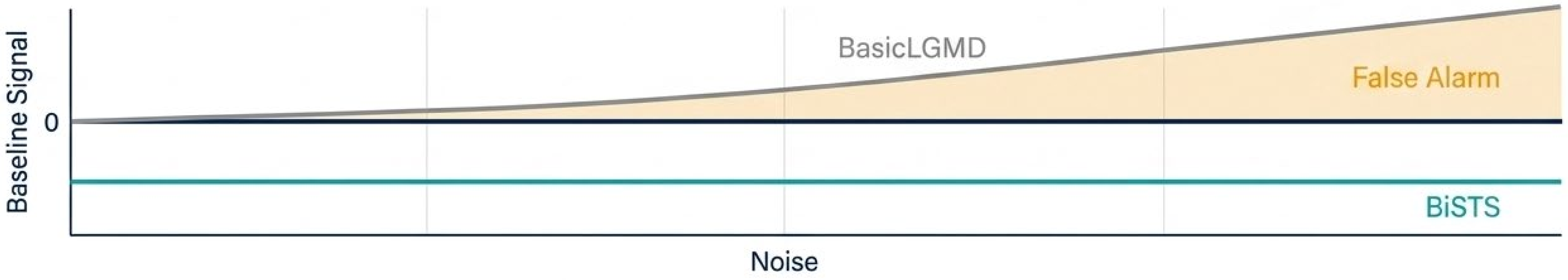
$$\mathcal{N}(x,y,t) \sim \mathcal{N}(0,\sigma_0^2)$$

- BasicLGMD yields $E(S) > 0$ (False Positive Escalation).
- BiSTS yields $E(S) < 0$ (Guaranteed Suppression).

## Salt & Pepper Noise (Theorem B.4)

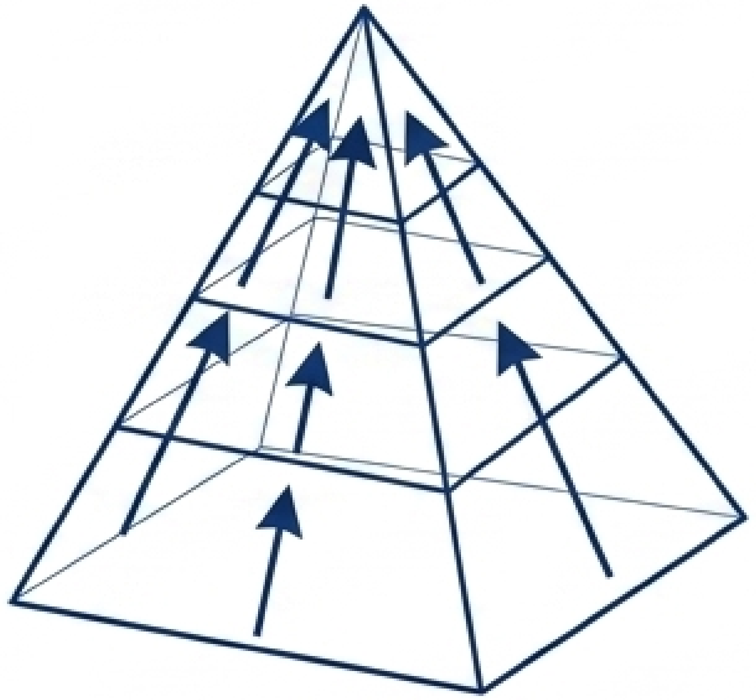
$$\text{Impulsive noise density } d \text{ in } (0,1]$$

- BasicLGMD fails catastrophically when $d > d^*$ in (0.11, 0.22).
- BiSTS consistently maintains $E(S) < 0$.



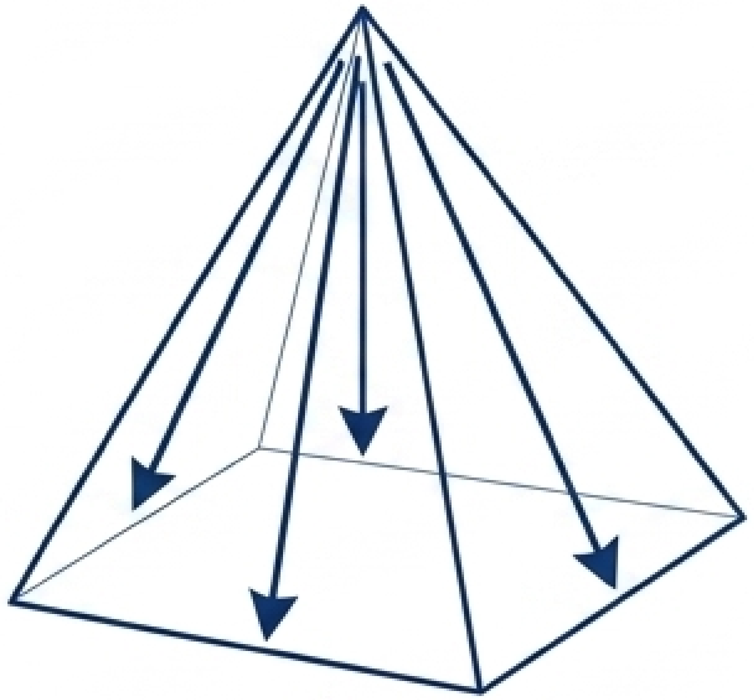
# Phase 2: Long-Term Spatiotemporal Coupling (LTSC)

Constructing hierarchical pyramids to correlate abstract spatial scale with historical temporal depth.



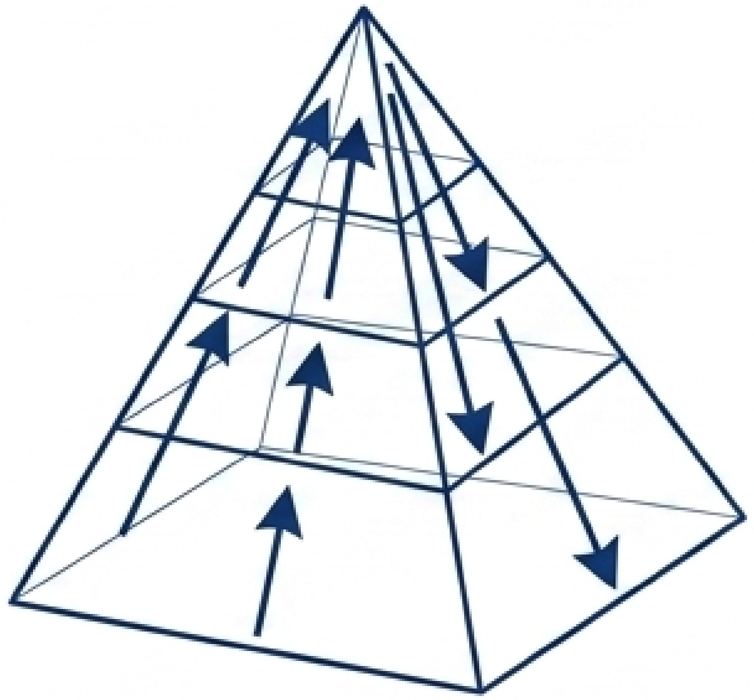
Spatial Pyramid ($P_S$)

Captures multiscale motion features and semantic context.



Temporal Pyramid ($P_T$)

Aggregates historical motion states.



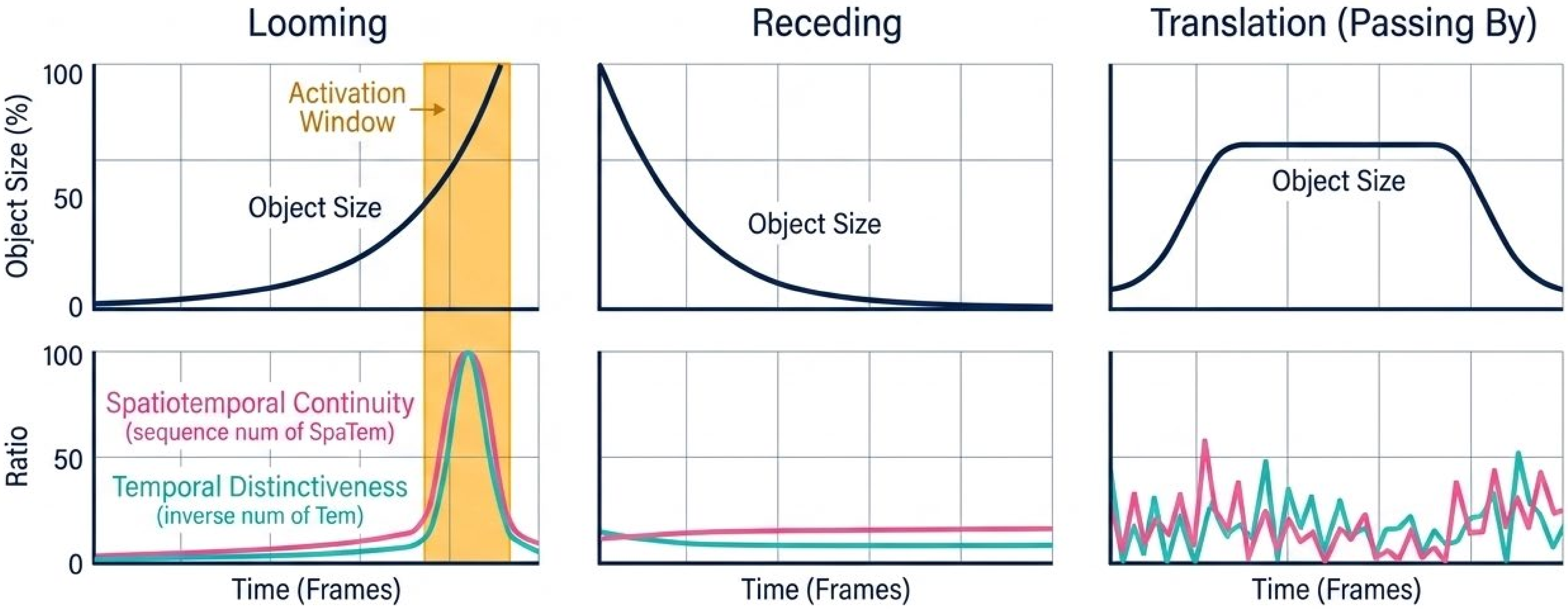
Spatiotemporal Pyramid

$$P_{ST} = P_S \cdot P_T$$

True looming motion is unique: its high-level spatial expansion aligns perfectly with its long-term temporal history. LTSC mathematically identifies this exact alignment.

# Isolating the unique signature of an approaching threat

LTSC tracks temporal distinctiveness and spatiotemporal continuity to suppress non-collision motion.

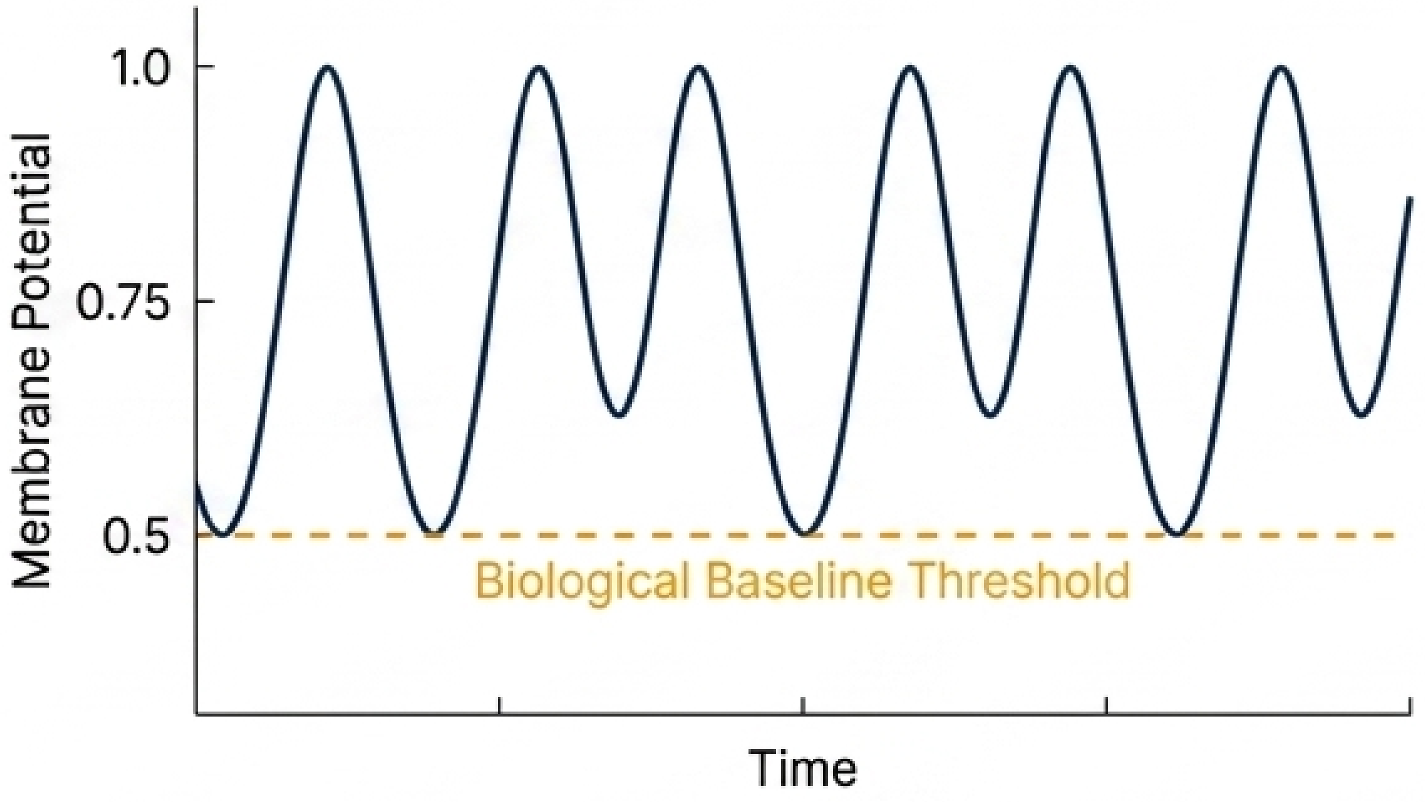



Traditional models trigger based on the sheer volume of moving pixels. BiSTS requires continuous spatiotemporal validation, effectively eliminating false alarms from lateral traffic.

# The evaluation flaw: Redefining benchmark metrics

Traditional Precision-Recall and AUC fail to capture the continuous, threshold-independent dynamics of biological models.

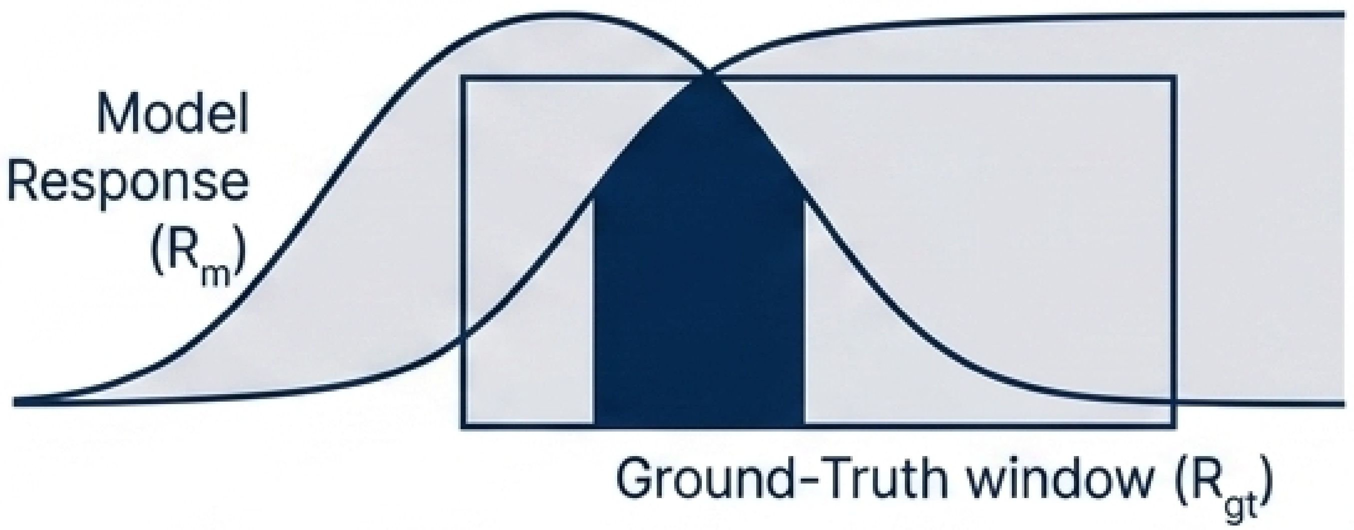
## The Flaw




Standard metrics rely on a fixed, often unattainable activation threshold. Because LGMD outputs are sigmoidal (bounded 0.5 - 1.0), baseline thresholds are biologically anchored at 0.5. Artificial thresholds are meaningless.

## The Solution - SA Metric

$$SA = \frac{R_m \cap R_{gt}}{R_m \cup R_{gt}}$$




Separability (SA) evaluates the entire physically meaningful response range, heavily penalizing false alarms outside the collision window.

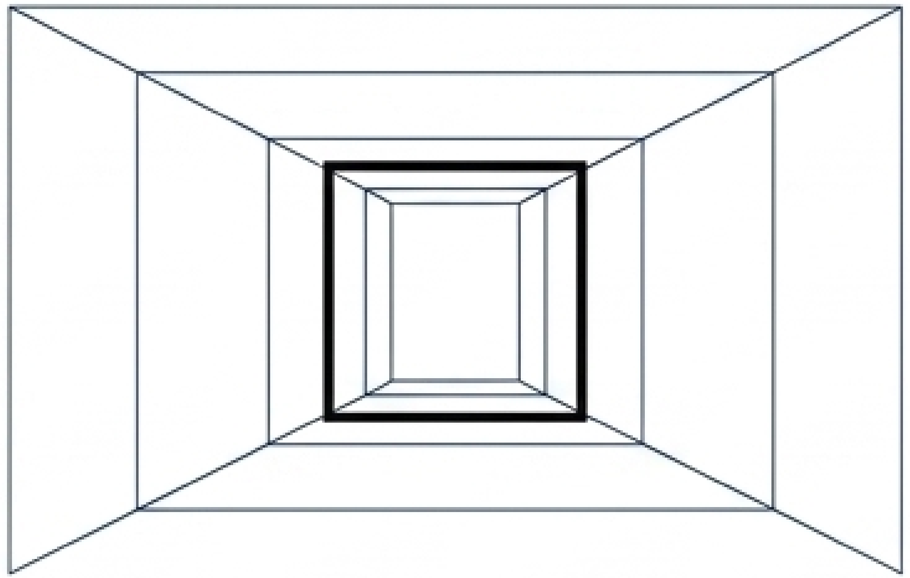
# A comprehensive benchmark for dynamic motion perception

Validating across 3,447 sequences spanning pure simulation, synthetic blending, and real-world edge scenarios

Pillar 1
**Simulation** (Basic Test)

## 16 sequences

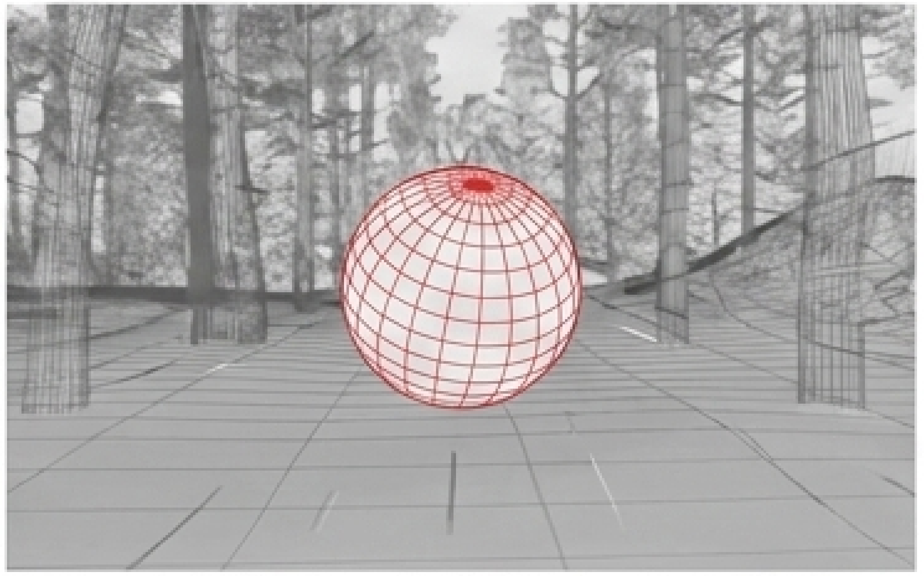
Pure kinematics. Tests ideal looming vs. receding/translation.



Pillar 2
**Synthesis** (Controlled Backgrounds)

## 3,208 sequences

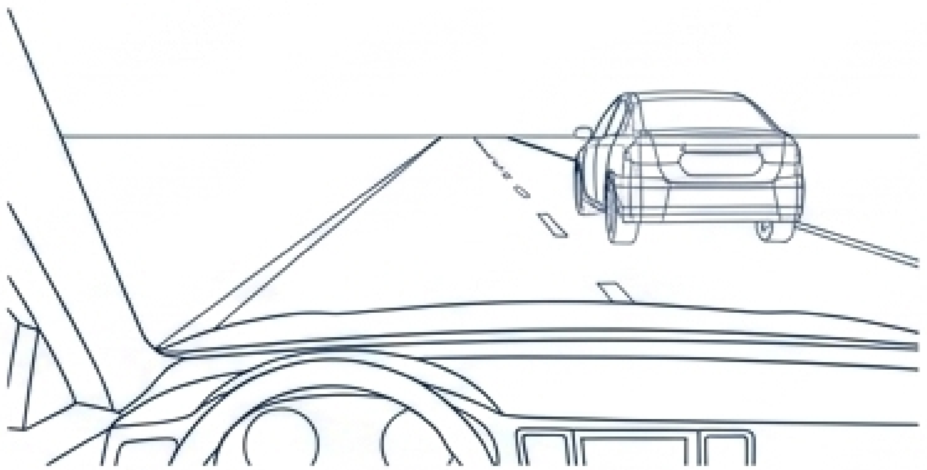
Looming objects injected into actual moving dashboard footage and complex woodlands. Tests figure-ground separation.



Pillar 3
**Real-World** (Physical Constraints)

## 223 sequences

Actual dashboard footage of physical vehicular collisions and high-speed non-clash driving. Tests unconstrained physical reliability.



# State-of-the-Art quantitative performance

BiSTS achieves a 51.8% improvement in Weighted Average Separability over the best existing bio-inspired baselines.

| Model | Basic Simulation (SA) | Complex Synthesis (SA) | Real-World Driving (SA) | Weighted Avg SA |
|---|---|---|---|---|
| BasicLGMD | 63.4 | 14.2 | 7.9 | 14.3 |
| MC-LGMD | 22.5 | 8.2 | 7.8 | 9.8 |
| pLGMD | 63.0 | 15.3 | 8.9 | 16.6 |
| **BiSTS (Ours)** | **74.4** | **16.4** | **13.4** | **25.2** |

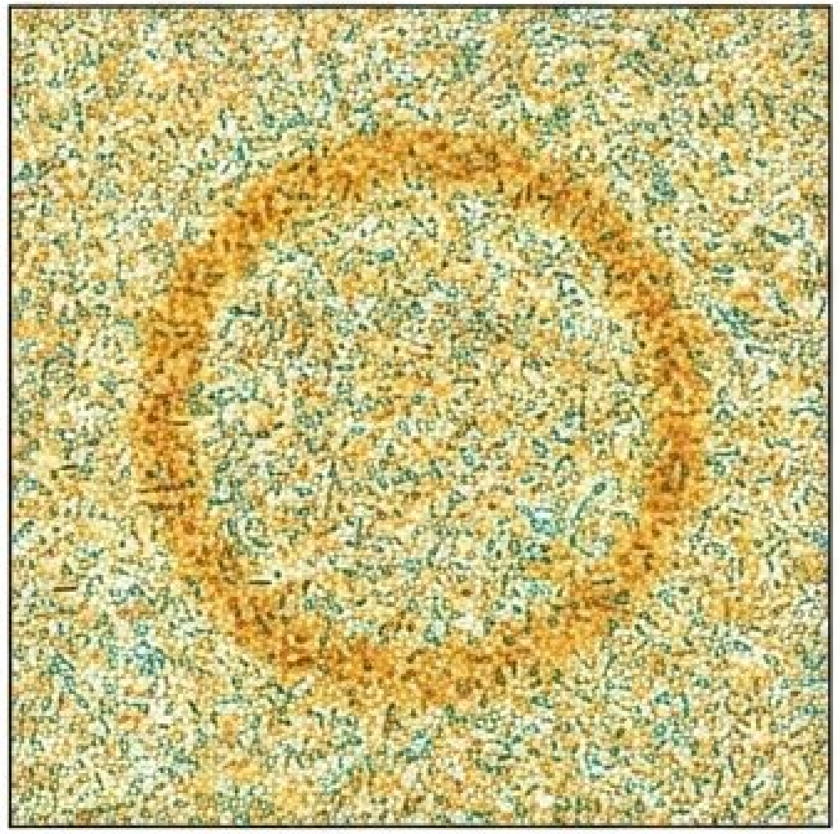
**+51.8% Improvement**

Even in highly cluttered driving scenarios where baseline models fail completely (SA < 9.0), BiSTS maintains functional target isolation.

# Real-world visual evidence of noise immunity

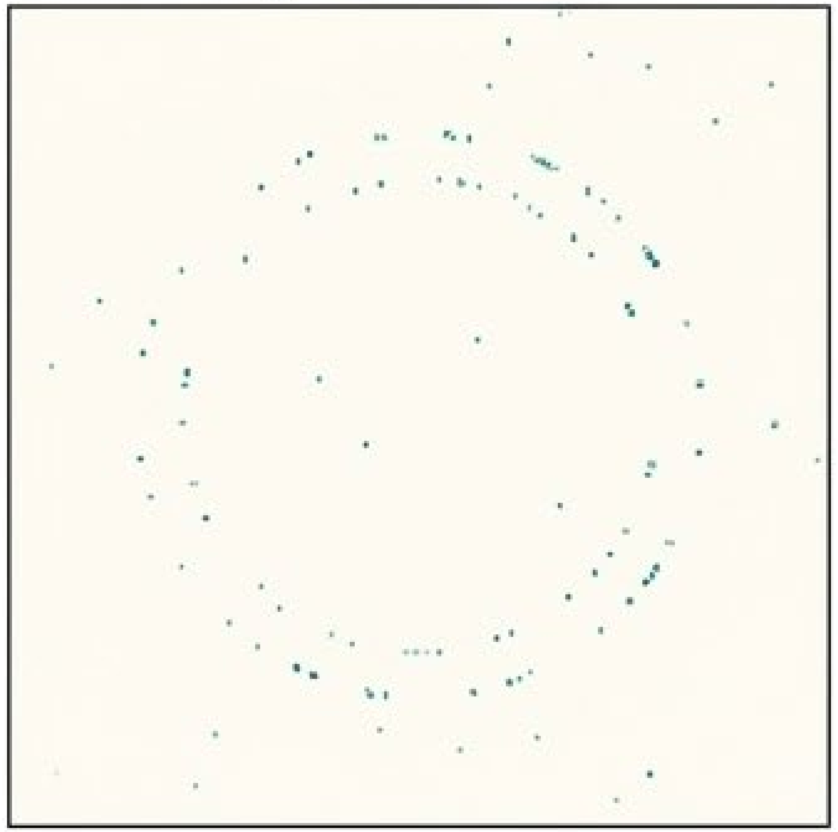
Maintaining target separability even under extreme 10% Salt & Pepper interference.

BasicLGMD



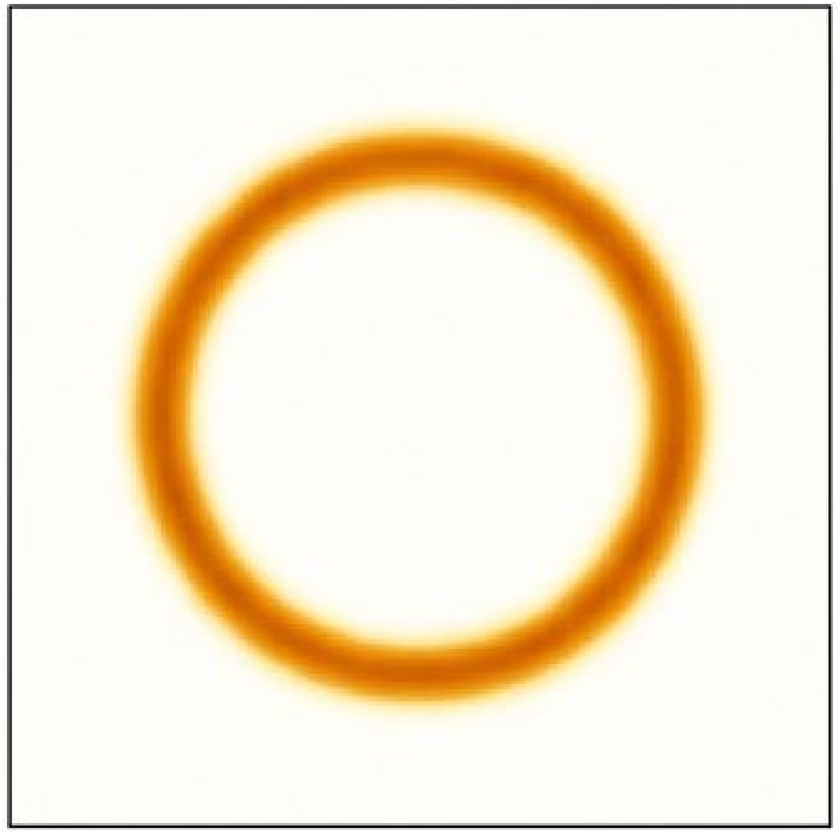
Peak Signal-to-Noise Ratio drops to 17.15 dB

pLGMD



PSNR 18.03 dB

BiSTS (Ours)

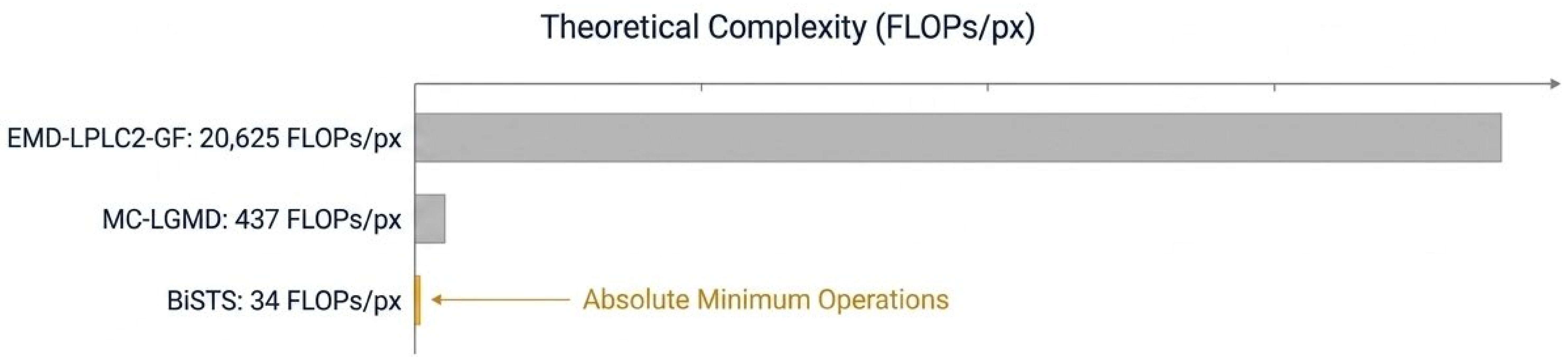


PSNR 23.58 dB

While probabilistic models discard essential structural cues to reduce noise, the BiSTS STSA module specifically filters the interference while preserving the target's physical edge.

# Physical edge deployment and SWaP efficiency

Breaking the accuracy-efficiency bottleneck on resource-constrained hardware (NVIDIA Jetson AGX Xavier).




Physical Edge Efficiency

**Throughput: 158.5 FPS**

Far exceeding the >30 FPS real-time autonomous driving requirement.

**Power Efficiency: 270.2 FPS/Watt**

Tested in constrained 15W mode.

High-fidelity collision avoidance utilizing a fraction of the computational and energy envelope required by traditional deep learning.

# Architectural ablation & contrast resilience

LTSC explicitly replaces Contrast Normalization (CN), while STSA supersedes the traditional Grouping (G) layer.

## Contrast Tests

| | Positive ($C>0$) | Negative ($C<0$) |
|---|---|---|
| BasicLGMD | 66.5 | 66.5 |
| BiSTS | **78.4** ($C>0$) | **79.1** ($C<0$) |

## Component Replacement Insight

- **Observation 1:** Removing STSA drops noise robustness from 44.1 to 38.0. Reintroducing a traditional ‘G’ layer yields 0% marginal improvement.
- **Observation 2:** Removing LTSC drops overall performance instantly. Adding explicit ‘CN’ pathways actually reduces synthetic benchmark performance.

**BiSTS structural mechanisms inherently subsume legacy filtering and normalization layers, radically simplifying the processing pipeline.**

# Summary of core contributions

## A new baseline for reliable, energy-efficient edge collision warning

### Mathematical Rigor

Engineered the **STSA** and **LTSC** modules to provide mathematically guaranteed, **biphasic noise suppression** and **contrast** *resilience*.

### Evaluation Standard

Established the first comprehensive 3.4k sequence **LGMD** benchmark and introduced the threshold-independent **Separability** $(SA)$ metric.

### Edge-Ready SOTA

Achieved an unmatched $51.8\%$ performance boost while operating at $158.5\ \mathrm{FPS}$ and $270.2\ \mathrm{FPS/W}$ on physical **15W edge GPUs**.

## Future directions in bio-inspired spatiotemporal processing

Bridging learning-free biological efficiency with data-driven adaptability.

# Transitioning from handcrafted biological parameters to trainable spatiotemporal architectures.

- Superseding static pooling pyramids with learnable convolutional layers.
- Automatically optimizing receptive fields for unstructured, highly variable environments.
- Preserving the ultra-low SWaP profile of the LGMD pathway while matching Deep Learning generalizability.

View the Benchmark
& Code Repository.