Boosting Open Set Recognition Performance through Modulated Representation Learning

Amit Kumar Kundu

Department of Electrical and Computer Engineering
Institute for Health Computing
University of Maryland, College Park
email: amit314@umd.edu

International Conference on Learning Representations (ICLR) 2026

The Open Set Problem

- **Reality:** Novel, unknown classes during inference[1].
- **Result:** Unreliable predictions, performance drops.

Figure: Closed set vs. Open set recognition scenario.

[1] Walter J Scheirer et al. (2012). "Toward open set recognition." In: *IEEE transactions on pattern analysis and machine intelligence* 35.7, pp. 1757–1772.

The Dual Objective of OSR

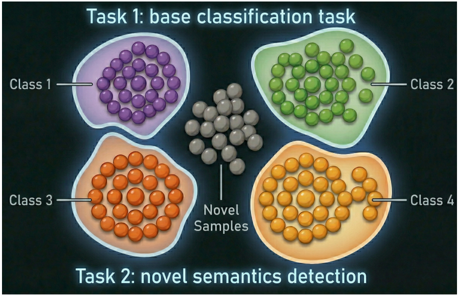
Two simultaneous tasks:

1. Base Task (Classification)

- **Goal:** Classify known classes. Needs **compact semantic clusters**.

2. OSR Task

- **Goal:** Flag unseen classes. Needs: **Strict boundaries for novelties**.



Challenges & Limitations in OSR

Core Difficulties

- **No auxiliary data. No info about novel semantics**
- **Feature overlap.**

Limitations of Prior Works

- **Burdening Base Task:** Heavy compute (generative models[a], mix-ups[b][c], or secondary models[d]).
- **Static Compromise:** Fixed temperatures in the loss function.

[*a*] Moon et al. (2022). “Difficulty-aware simulator for open set recognition.” In ECCV.
[*b*] Xu et al. (2023). “Contrastive open set recognition.” In: AAAI.
[*c*] Li et al. (2024). “All beings are equal in open set recognition.” In: AAAI.
[*d*] Zhou et al. (2024). “Contrastive learning based open-set recognition with unknown score.” In: Knowledge-Based Systems.

Requirement: Rich representations, dual-task satisfaction, **minimal overhead**.

Our Contributions

- **Novel Modulation:** Introduced Negative Cosine Schedule (NegCosSch) for temperature modulation.
- **Zero Overhead:** Plug-and-play into existing losses (CE, SupCon, ARPL).
- **Performance Boosts** of SOTA methods
- **Scalability:** Gains increase with class complexity.

Temperature Dynamics: The Low τ Extreme

Gradient w.r.t negative l_j

$$\frac{\partial L_{SupCon}}{\partial l_j} = \frac{1}{\tau}[\text{softmax}_{a \in I \setminus \{i\}}(\text{sim}(l_i, l_a)/\tau)]_j \times \frac{\partial \text{sim}(l_i, l_j)}{\partial l_j}$$

How does τ shape the manifold?

Low $\tau \to 0$ (Instance Focus[a]):

- Gradient heavily dominates *hard* negatives.
- **Benefit:** Forces sharp boundaries, pushing unknowns far away.
- **Flaw:** fails to form compact semantic clusters.

[a] Wang et al. (2021)."understanding the behaviour of contrastive loss." In: CVPR

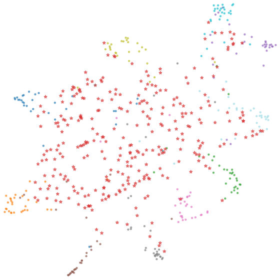


Figure: Low $\tau = 0.5$: Sharp boundaries, loose clusters.

Temperature Dynamics: The High τ Extreme

Gradient w.r.t negative l_j

$$\frac{\partial L_{SupCon}}{\partial l_j} = \frac{1}{\tau}[\text{softmax}_{a \in I \setminus \{i\}}(\text{sim}(l_i, l_a)/\tau)]_j \times \frac{\partial \text{sim}(l_i, l_j)}{\partial l_j}$$

How does τ shape the manifold?

High $\tau \to \infty$ (Class Focus):

- Gradient shift towards *easy* negatives.
- **Benefit:** Creates compact semantic clusters.
- **Flaw:** Loses discriminative power; unknowns easily overlap with knowns.

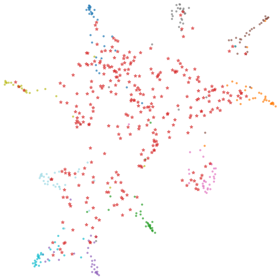


Figure: High $\tau = 2.0$: Compact clusters, overlapping

Proposed Solution: Temperature Modulation

The Ultimate Rationale

Both features required!!! Scheduling traverses this spectrum!!!

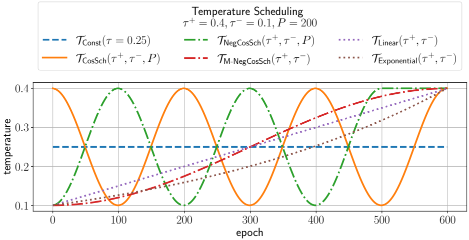


Figure: Different temperature schedules.

Proposed: Negative Cosine Scheduling

$$\mathcal{T}_{\text{NegCosSch}}(e; \tau^+, \tau^-, P, k) = \begin{cases} \tau^- + \frac{1}{2}(\tau^+ - \tau^-)(1 + \cos(\frac{2\pi e}{P} - \pi)), & \text{if } e \leq E - \frac{P}{2} \\ \tau^+, & \text{elsewhere} \end{cases}$$

NegCosSch: Manifold Evolution during a Cycle

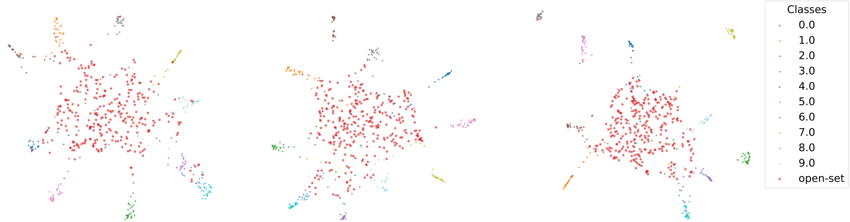
Transforming the feature space from instance-focus to cluster-focus:

Start: Discrimination (Low τ)

- **Initial Focus:** Hard negatives.
- **Manifold:** Carves out sharp boundaries around knowns.
- **Goal:** Establish the "exclusion zone" for unknowns.

End: Refinement (High τ)

- **Final Focus:** Easy negatives.
- **Manifold:** smooths into compact semantic clusters.
- **Goal:** Maximize base accuracy while preserving isolation.



(a) Start (Low τ) (b) Mid-cycle (c) End (High τ)

Result: A perfectly optimized space that satisfies both task requirements!!!

Performance Highlight

Table: Performance on CUB ('Easy / Hard' splits)

Methods	Acc. (%)	AUROC (%)	OSCR (%)
CE Constant	84.43	83.55 / 74.98	70.49 / 63.34
CE + **NegCosSch(ours)**	**86.12**	**86.79** / **78.08**	**74.70** / **67.30**
SupCon Constant	83.43	86.94 / 73.95	72.42 / 61.66
SupCon + **NegCosSch(ours)**	**85.30**	**88.14** / **75.81**	**75.09** / **64.72**

Key Result

NegCosSch improves all baselines. AUROC boosts up to **+3.3%** and OSCR up to **+4.4%**. No overhead.

Ablation Study: Temperature Schedules

Table: Comparison of different schedules on CUB.

Schedule	Accuracy (%)	AUROC (%)	OSCR (%)
Constant (Baseline)	84.43	83.55 / 74.98	70.49 / 63.34
Linear decrease	81.64	79.86 / 71.75	65.15 / 58.59
CosSch	84.63	84.5 / 74.24	71.51 / 62.93
Logarithmic increase	85.15	84.91 / 76.07	72.25 / 64.82
Exponential increase (ours)	**86.12**	**86.65** / **78.05**	**74.64** / <u>**67.35**</u>
Linear increase (ours)	**86.22**	86.54 / **78.01**	74.58 / **67.32**
P-NegCosSch (ours)	<u>**86.3**</u>	<u>**86.85**</u> / 77.6	<u>**74.89**</u> / 67.01
M-NegCosSch (ours)	**86.12**	**86.79** / <u>**78.08**</u>	**74.7** / **67.3**

Findings

- **Winner:** NegCosSch wins 12 out of 18 cases.

Robustness and Key Findings

- **Geometric Advantage:** Forces unknowns further from known prototypes via lower intra-class scatter and higher margin.
- **Universal Gains:** Improves baselines: **CE, SupCon, ARPL, Prototypical loss and BackMix**.
- **Backbone Agnostic:** Consistent AUROC/OSCR gains across **VGG, ResNet, and ViT**.
- **Metric Invariant:** transfers across **Energy, ODIN, and Max-prob** scoring rules.
- **LS Resilience:** Boosts performance even when Label Smoothing (LS) degrades the static baseline.
- **Scalability:** Improvement gap **widens** with class complexity.

Summary

- **Theory:** Fixed temperature forces sub-optimal trade-offs.
- **Solution:** NegCosSch sweeps τ (Low $\rightarrow$ High). Decouples objectives.
- **Efficiency:** Zero computational overhead. Plug-and-play.
- **Results:** Gains on SOTA.

Thank You!

Questions?

Support

- University of Maryland Institute for Health Computing (UM-IHC),
- University of Maryland Institute for Advanced Computer Studies (UMIACS),
- ICLR 2026