William Yang
4th year PhD
williamyang@princeton.edu

Princeton Visual AI Lab

PRINCETON UNIVERSITY

Xindi Wu
3rd year PhD
xindiw@princeton.edu

Demystifying Computer Vision Systems

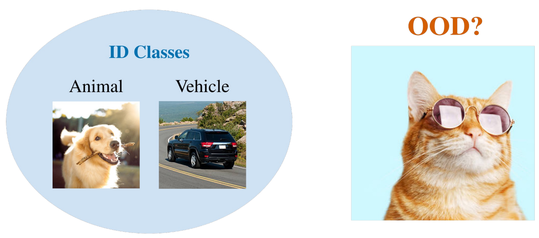
ImageNet-OOD: Deciphering Modern Out-of-Distribution Detection Algorithms

William Yang*, Byron Zhang*, Olga Russakovsky ICLR, 2024

To study impact of semantic shift without the influence of covariate shift in new-class detection, we construct **ImageNet-OOD** from ImageNet-21K

Semantic Ambiguity

Visual Ambiguity

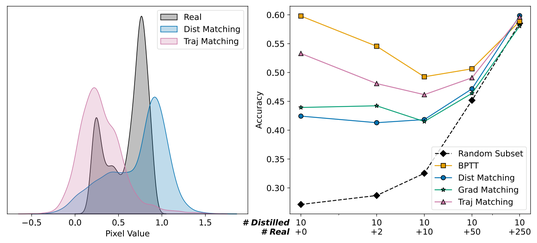


Class Ambiguity w.r.t the Decision Boundary

What is Dataset Distillation Learning

William Yang, Ye Zhu, Zhiwei Deng, Olga Russakovsky ICML, 2024

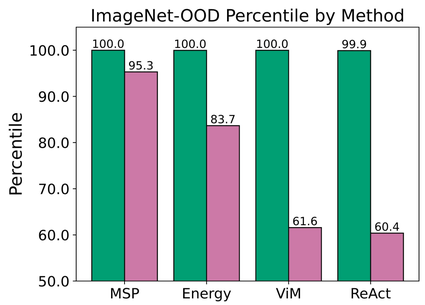
Dataset distillation learns a compact set of synthetic data that retains essential information from the original dataset, but why this is possible and what do they represent remains unclear.

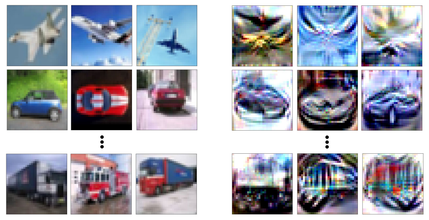


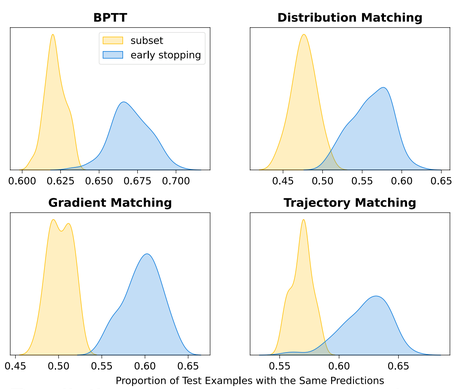
The distilled data lies outside of the real data manifold and is sensitive during training.

What are modern OOD detectors detecting?

One qualitative example on how covariate shift affects OOD detection performance







Predictions from models trained on distilled data are more similar to models that were early-stopped on the full data.

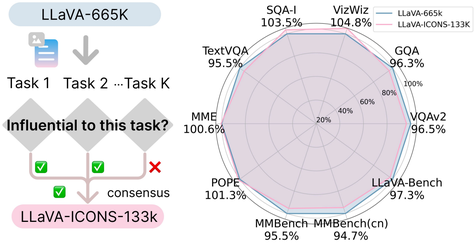
Data for Efficient Multimodal Machine Learning

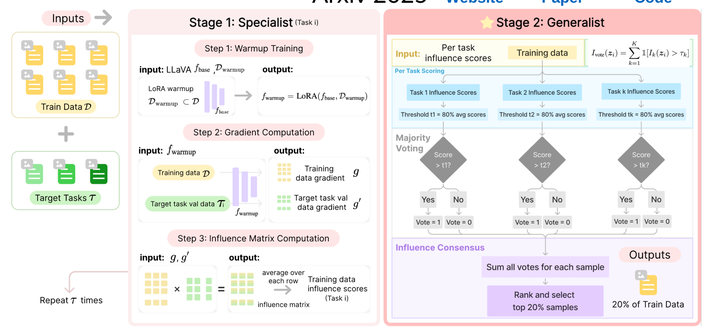
ICONS: Influence Consensus for Vision-Language Data Selection

Xindi Wu, Mengzhou Xia, Rulin Shao, Zhiwei Deng, Pang Wei Koh, Olga Russakovsky

Arxiv 2025 Website Paper Code Dataset

Can we identify a compact subset of training data that preserves model capabilities while enabling faster experimentation?





Key results:

A 20% subset of LLaVA-665K and a 20% subset of Cambrian-7M achieve 98.6% and 98.8% of the rel. obtained using the full datasets. LLaVA 60% subset achieves>100% rel.
Demonstrated strong transferability across unseen tasks and scalability to larger models.

ConceptMix: A Compositional Image Generation Benchmark with Controllable Difficulty

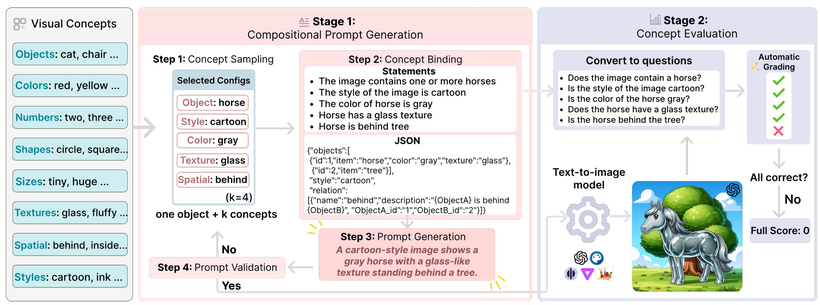
Xindi Wu*, Dingli Yu*, Yangsibo Huang*, Olga Russakovsky, Sanjeev Arora

Website Paper Code

NeurIPS, 2024

How to automatically generate prompts that mix diverse visual concepts?

How to automatically grade results based on (prompt, generated image)?



Key results:

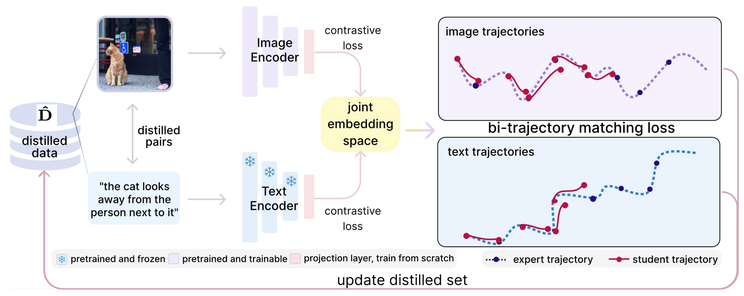
Color and style are the easiest while spatial, size, and shape are challenging.
We observe a consistent performance drop as K increases with the leading model DALLE 3, struggles at k=5.

Why are models bad at $k \geq 3$?

- Limited exposure to complex concept combinations.
- Disparate concept representation.

Vision-Language Dataset Distillation

Xindi Wu, Byron Zhang, Zhiwei Deng, Olga Russakovsky



- **Bi-trajectory matching**: Separately considers two trajectories to capture complex vision-text interactions via contrastive loss.

How can we distill the most critical information from vision-language datasets?

Website arXiv Code

TMLR, 2024

Distilled examples:

Key results:

Our method doubles the performance of coreset selection approaches with an order of magnitude fewer distilled pairs.