MultiOOD: Scaling Out-of-Distribution Detection for Multiple Modalities

Hao Dong, Yue Zhao, Eleni Chatzi, Olga Fink

NeurIPS 2024

ETH zürich

EPFL

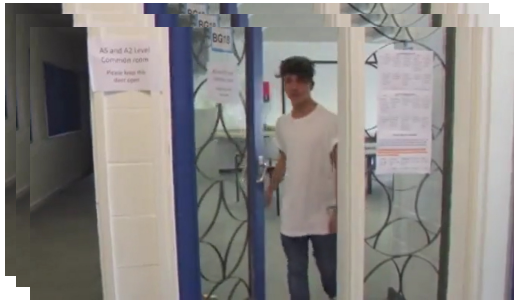
Motivation

- Real world is dynamic

Motivation

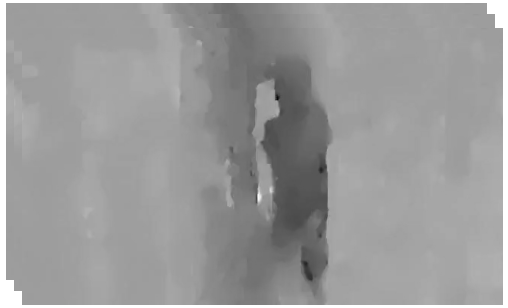
- Real world is multimodal

RGB images



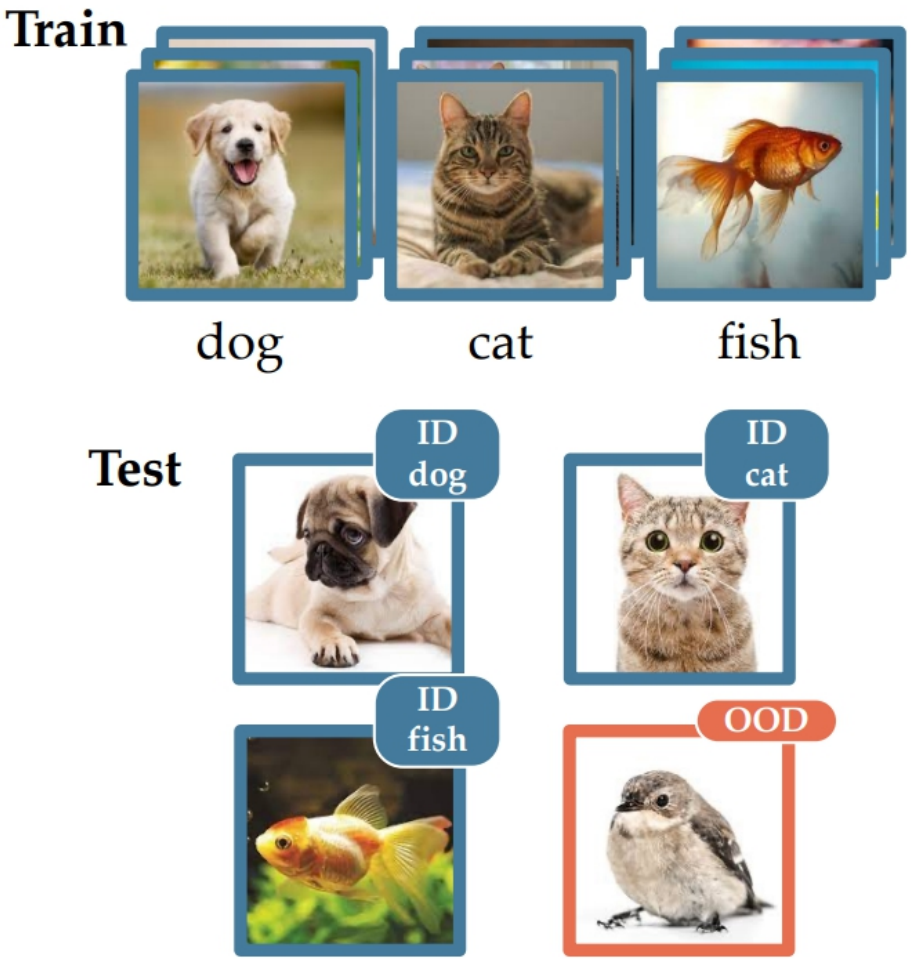
Video

LiDAR

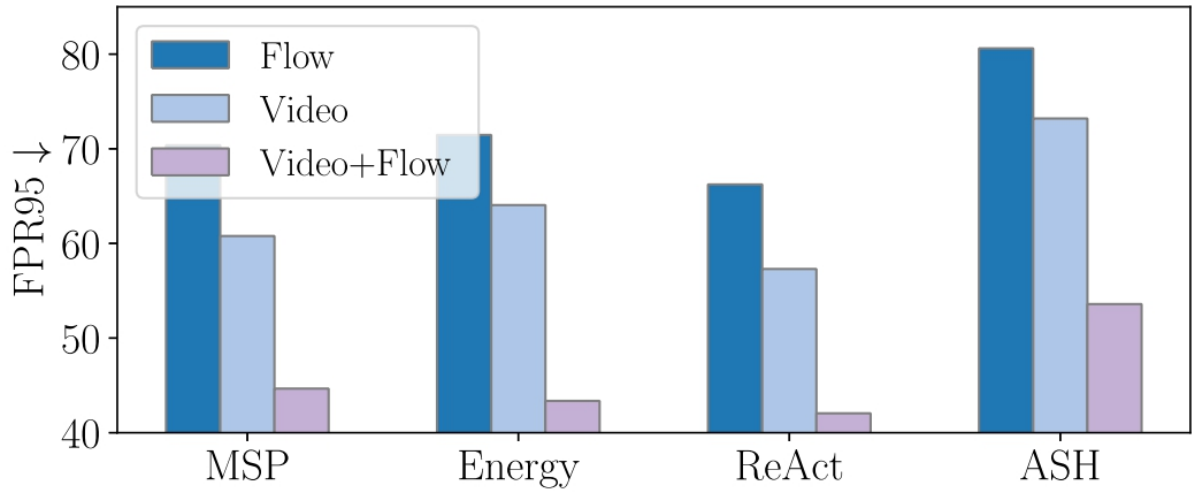


Optical Flow

Out-of-Distribution Detection

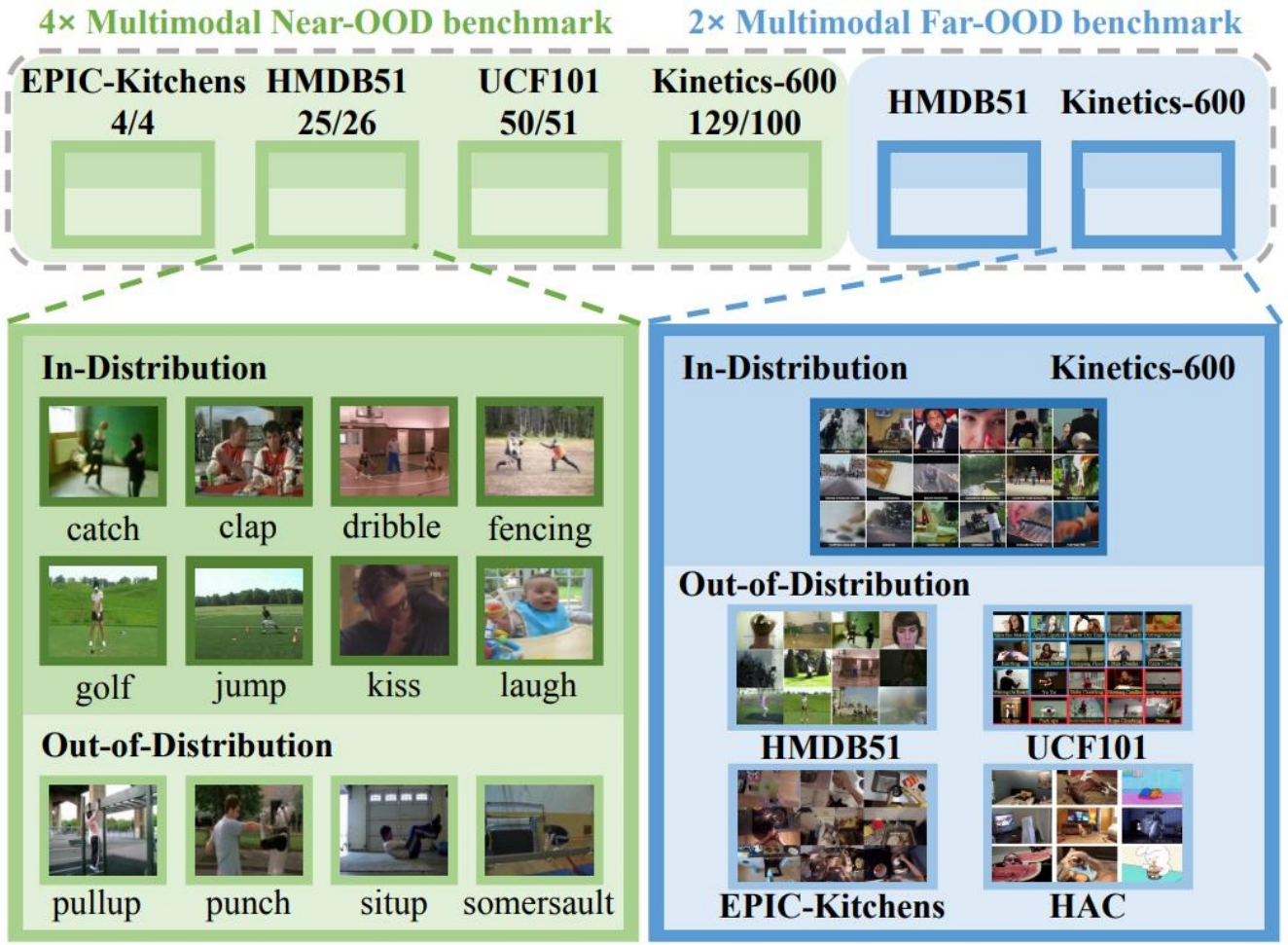


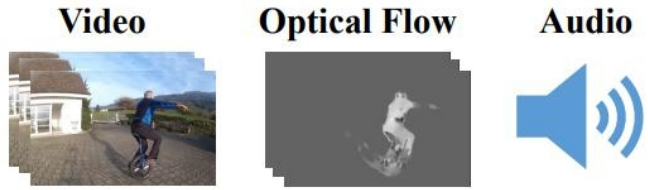
Why Multimodal

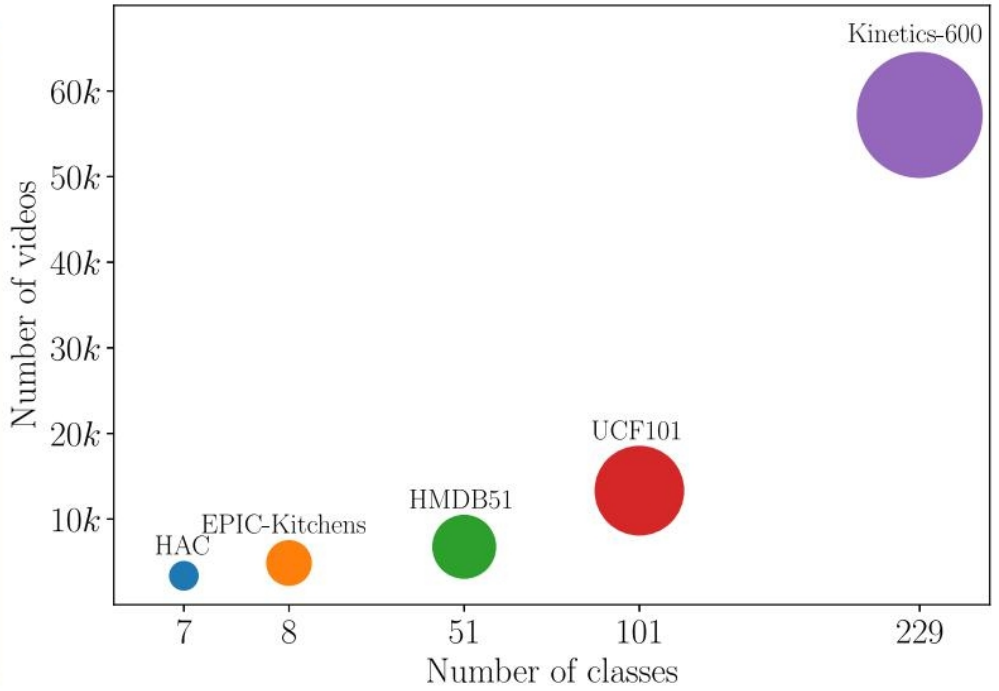


MultiOOD Benchmark

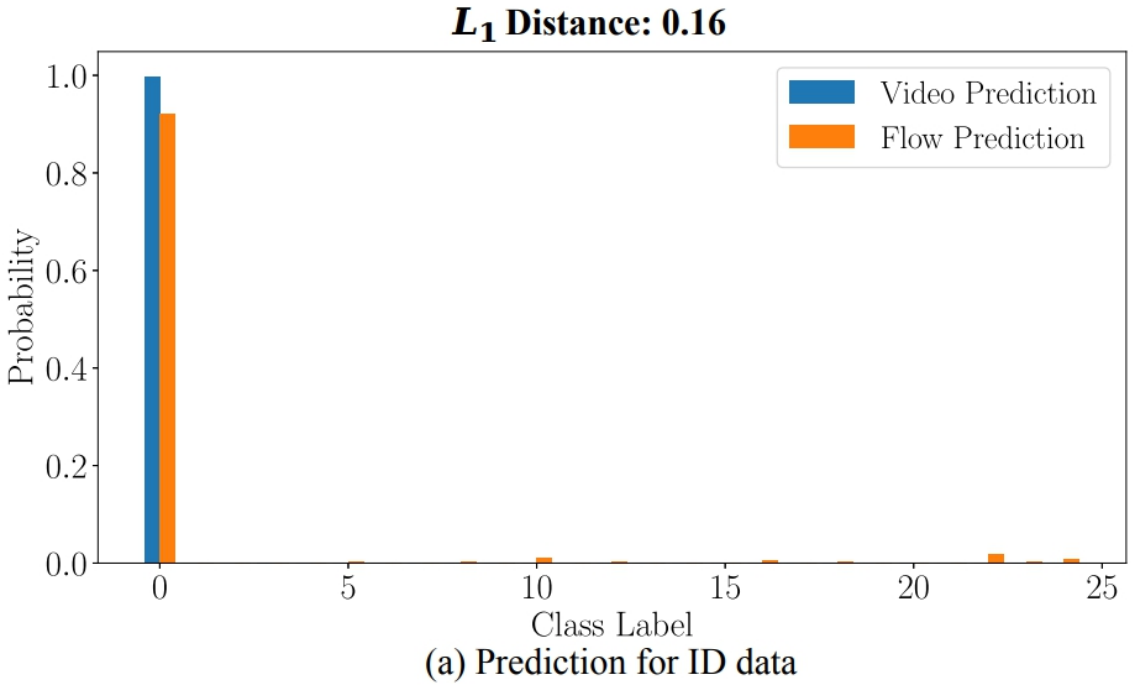
- The ***first*** benchmark for Multimodal OOD Detection



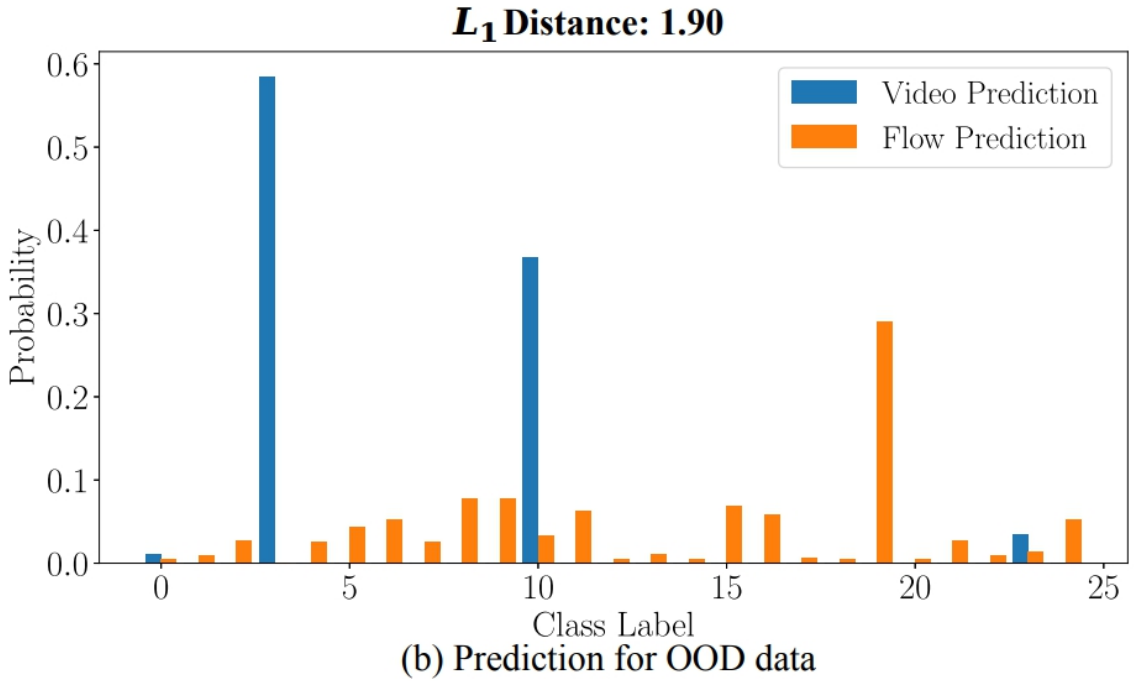




Modality Prediction Discrepancy

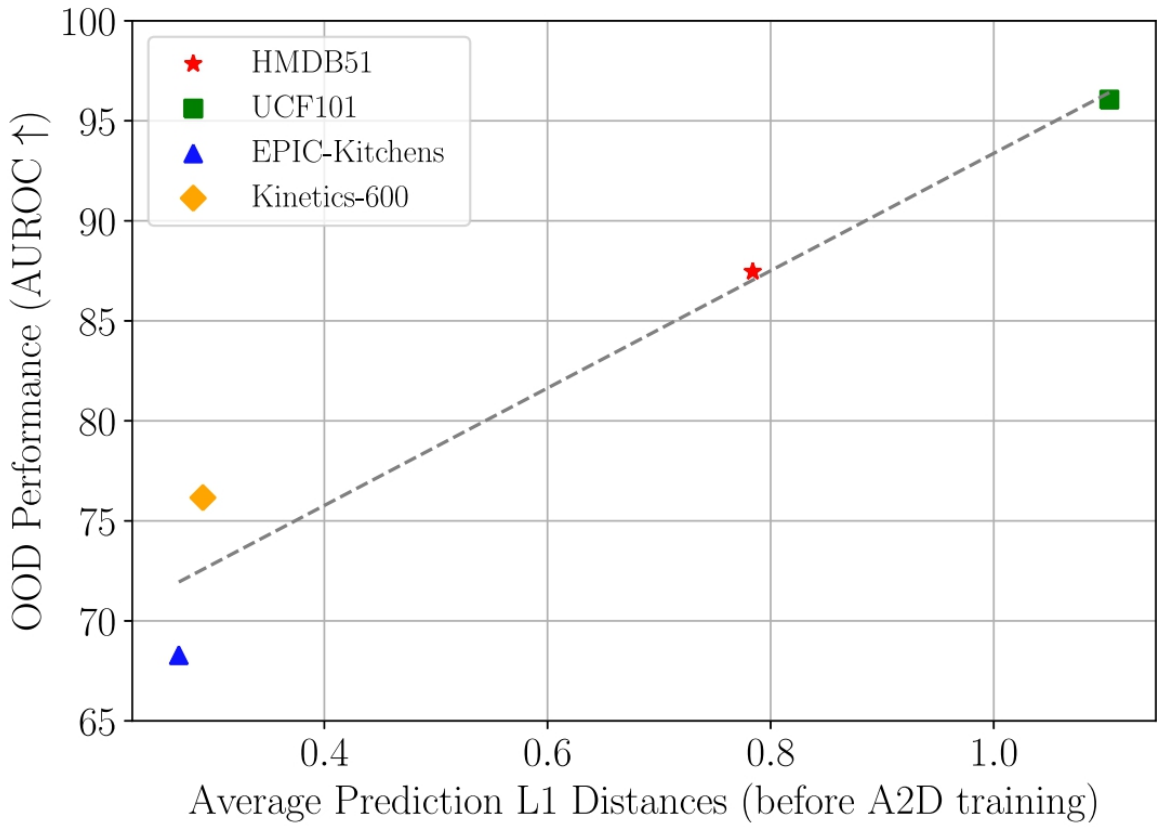


(a) Prediction for ID data

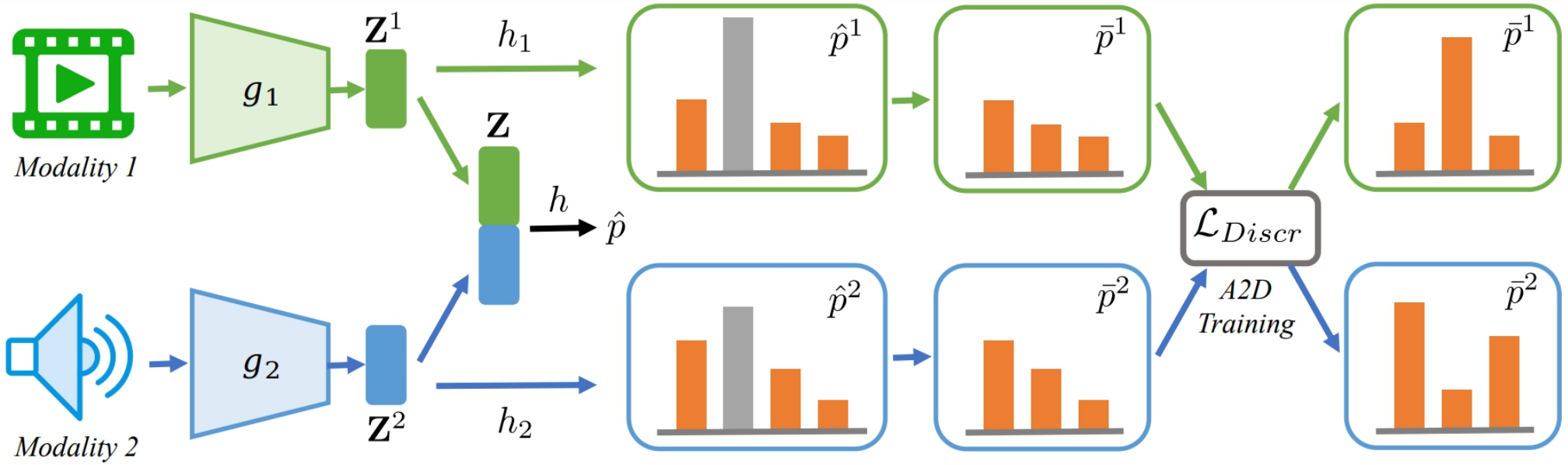


(b) Prediction for OOD data

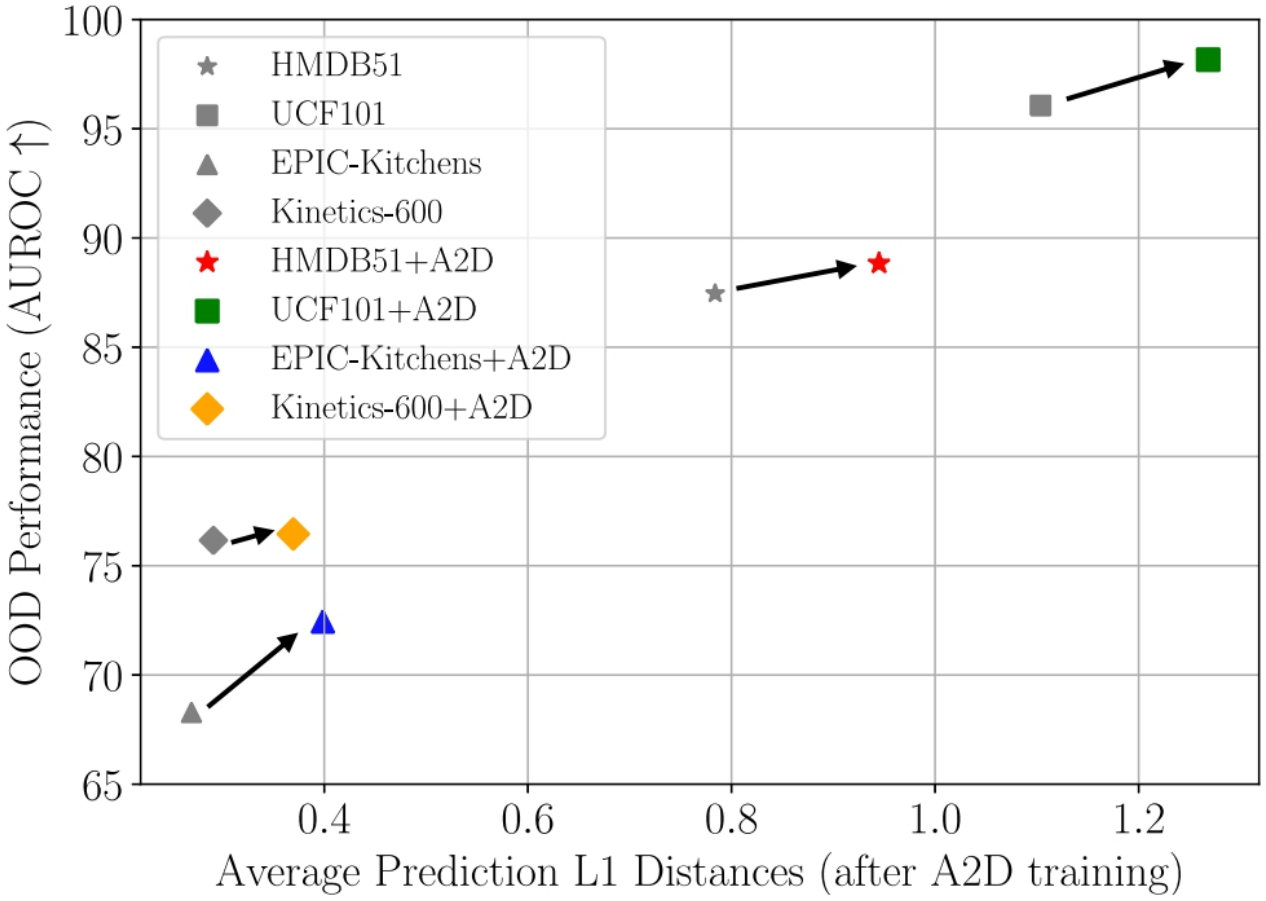
Modality Prediction Discrepancy



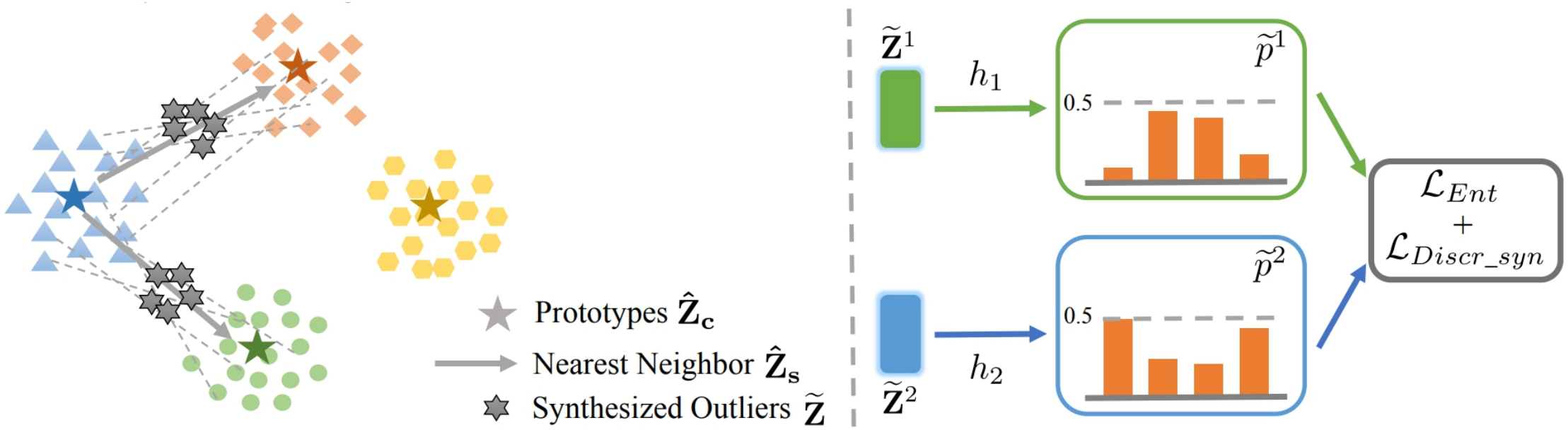
Agree-to-Disagree Algorithm



Agree-to-Disagree Algorithm



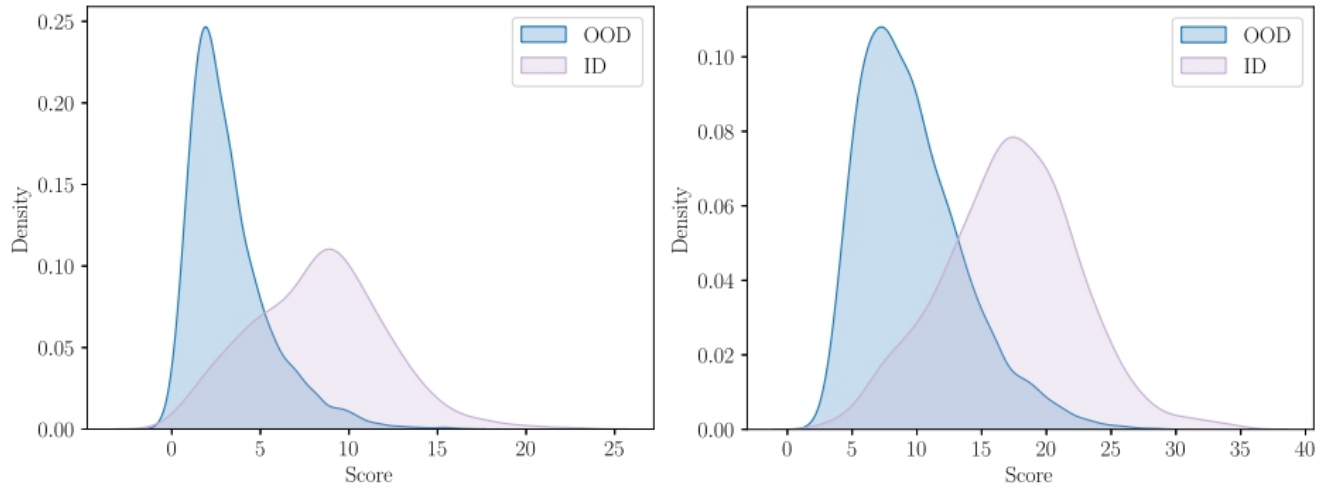
Nearest Neighbor Prototype-based Mixup



Multimodal OOD Detection with HMDB51 as ID

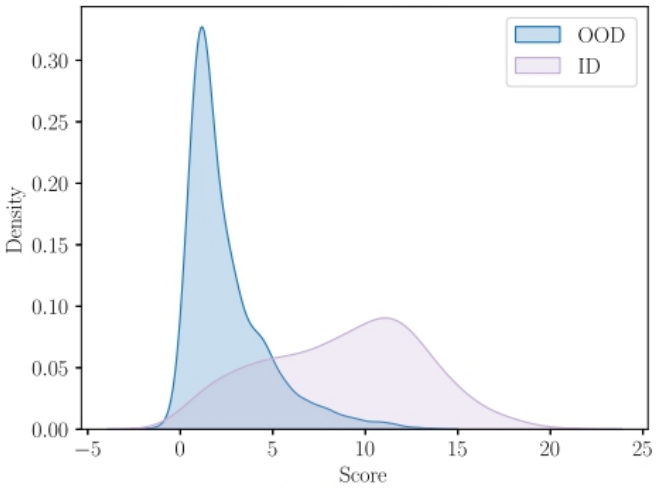
Methods	OOD Datasets										ID ACC ↑
	Kinetics-600		UCF101		EPIC-Kitchens		HAC		Average		
	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑	
Without A2D Training											
Energy	32.95	92.48	44.93	87.95	8.10	97.70	32.95	92.28	29.73	92.60	87.23
ASH	51.20	87.81	53.93	84.22	19.95	95.92	42.99	90.23	42.02	89.55	86.20
GEN	41.51	90.34	46.18	87.91	8.21	98.26	38.31	91.28	33.55	91.95	87.23
KNN	22.69	95.01	39.34	89.28	9.92	97.92	20.75	96.02	23.18	94.56	87.23
VIM	13.68	97.01	33.87	91.45	5.93	98.15	13.45	97.12	16.73	95.93	87.23
With A2D Training and NP-Mix Outlier Synthesis											
Energy++	$24.52_{-8.43}$	$93.96_{+1.48}$	$36.49_{-8.44}$	$89.67_{+1.72}$	$6.96_{-1.14}$	$97.53_{-0.17}$	$22.92_{-10.14}$	$94.41_{+2.13}$	$22.72_{-7.01}$	$93.89_{+1.29}$	86.89
ASH++	$27.82_{-23.38}$	$93.17_{+5.36}$	$38.43_{-15.50}$	$89.52_{+5.30}$	$6.84_{-13.11}$	$98.23_{+2.31}$	$23.03_{-19.96}$	$94.45_{+4.22}$	$24.03_{-17.99}$	$93.84_{+4.29}$	86.20
GEN++	$25.66_{-15.85}$	$93.50_{+3.16}$	$37.40_{-8.78}$	$91.19_{+3.28}$	$5.25_{-2.96}$	$98.98_{+0.72}$	$24.63_{-13.68}$	$94.28_{+3.00}$	$23.24_{-10.31}$	$94.49_{+2.54}$	86.89
KNN++	$15.05_{-7.64}$	$96.96_{+1.95}$	$33.06_{-6.28}$	$91.92_{+2.64}$	$5.47_{-4.45}$	$98.97_{+1.05}$	$13.45_{-7.30}$	$97.25_{+1.23}$	$16.76_{-6.42}$	$96.28_{+1.72}$	86.89
VIM++	$9.24_{-4.44}$	$98.04_{+1.03}$	$26.45_{-7.42}$	$92.34_{+0.89}$	$5.36_{-0.57}$	$98.09_{-0.06}$	$6.04_{-7.41}$	$98.56_{+1.44}$	$11.77_{-4.96}$	$96.76_{+0.83}$	86.89

Score Distributions

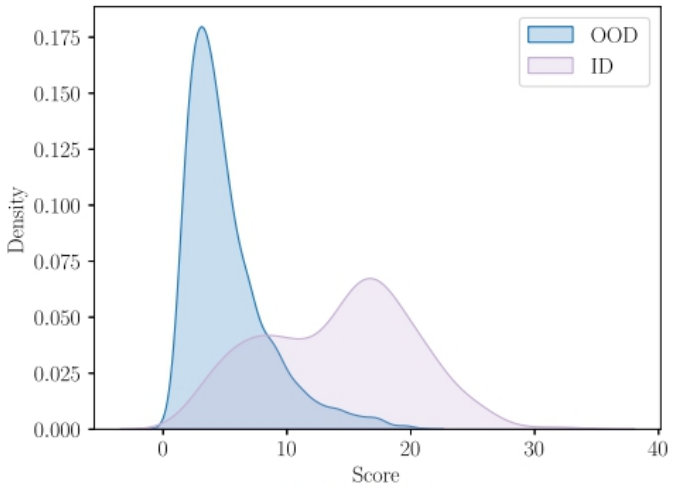


(a) Energy

(b) ASH



(e) Energy++



(f) ASH++

Conclusion

- The ***first*** benchmark for Multimodal OOD Detection
- A novel ***A2D*** training algorithm, inspired by the observation of the *Modality Prediction Discrepancy* phenomenon
- A new outlier synthesis algorithm ***NP-Mix***
- Extensive evaluations

Code and Benchmark: https://github.com/donghao51/MultiOOD