A NEW OOD DATASET FOR FINE-GRAINED CLASSIFICATION

指導教授：帥宏翰

專題生：電機大三 游棋翔

ABSTRACT

This paper introduces a novel Out-of-Distribution (OOD) dataset tailored for fine-grained classification challenges in machine learning. Addressing the **shortcomings of existing OOD datasets**, this work focuses on specificity and clarity, especially in domains with varied objects. By meticulously comparing class names across in-distribution (ID) and OOD datasets, we eliminate duplications, thereby enhancing the dataset's uniqueness. The result is **a more precise and reliable dataset for identifying fine-grained differences in complex scenarios**, particularly in the categorization of cars, bird species, and butterflies.

BACKGROUND

Traditional OOD datasets, such as ImageNet, often suffer from a lack of precision due to their broad scope and incidental ID elements. Our dataset diverges from this approach, offering higher precision through detailed categorization. The main issue with broader datasets is **the compromise in OOD integrity caused by non-distinctive features**. Our dataset's design overcomes this by **providing clear, distinct categories**, ensuring an enhanced level of fine-grained classification that is critical for advanced machine learning models.

EX.

Known that acorn_woodpecker is not one of species in CUB_200_2011

↓ CLASSIFIER

Show that this picture doesn't belong to Cub_200_2011's species

METHOD

To create a distinct dataset for fine-grained classification, we analyzed class name similarities across **Birds, Butterflies, and Cars** domains. Using LLM, we first get the specific detail(Order, Family, Genus, Species) of each class, then Python we filtered out overlapping species names from ID and OOD datasets. For Birds, we refined **NABirds** and **Birdsnap**, reducing the number of classes and files. **Butterfly-200** and **ETHEC** datasets were similarly trimmed. The Car category was curated using **Stanford Cars** and **Compcars**, leading to a more concise set of classes and files.

Ex.

CUB / Birdsnap / NABird

Common bird names

↓ GPT3.5 Turbo

Order, Family, Genus, Species

↓

Remove bird species duplicated in Birdsnap and NABird with those in CUB, based on genus+species names.

RESULT & DISCUSSION

We introduce an enriched Out-of-Distribution (OOD) dataset and benchmark it against those used in the MixOE paper by Zhang et al., and Pramuditha Perera et al. Our evaluation, as shown in Table 1, and the hierachy of bird species in MIXOE's dataset reveals a previously limited diversity in class and image numbers. MixOE's **holdout class method**, which creates ID and OOD splits, inadvertently reduces this diversity—vital for fine-grained object recognition. Our dataset counters this by widening the range of classes and images across Birds, Butterflies, and Cars domains. The Bird domain, combining NABirds and Birdsnap, and the Butterfly and Car domains, sourced from Butterfly-200, ETHEC, Stanford Cars, and Compcars, are substantially enhanced post-filtering. The application of current methodologies to our new ood dataset (Table 2) yields improved results in comparison with existing OOD datasets. Our carefully curated dataset not only offers a richer variety but also have better result on demarcating ID from OOD datasets, which is vital for detailed classification analysis.

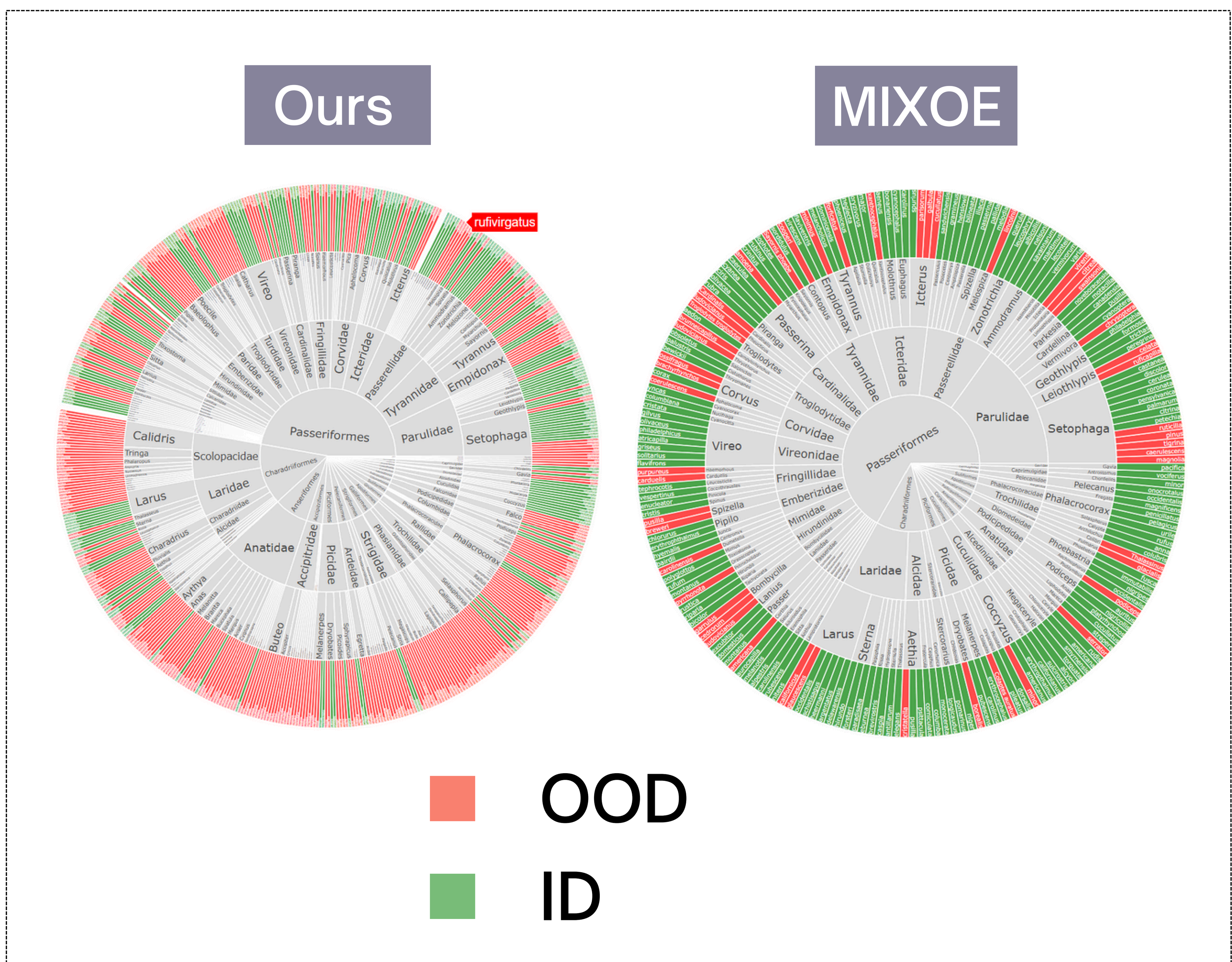

Table 1.

Methods	ID Dataset	ID # classes	ID # images	OOD Dataset	OOD # classes	OOD # images
MixOE	NABirds	200	7,300	NABirds	55	2,200
	Butterfly	150	7,000	Butterfly	50	3,600
	Stanford Cars	150	5,600	Stanford Cars	46	1,860
	Aircraft	90	5,400	Aircraft	12	330
Ours	CUB	200	11,788	NABirds	366	30,728
				Birdsnap	344	27,380
	Butterfly	200	25,279	ETHEC	647	39,681
	Stanford Cars	196	16,185	Comcar	1,693	134,318

Table 2.

		CUB-200-2011									
		Near-OOD				Far-OOD				Average	
		NAbirds		Birdsnap		ETHEC		Comcar			
Models	ID ACC	FPR95 ↓	AUROC ↑	FPR95 ↓	AUROC ↑	FPR95 ↓	AUROC ↑	FPR95 ↓	AUROC ↑	FPR95 ↓	AUROC ↑
MSP	82.06	55.70	83.15	53.14	84.59	6.94	98.31	7.61	98.20	30.85	91.06
TempScale	82.06	54.64	83.66	52.64	85.14	5.73	98.59	6.14	98.51	29.79	91.48
ODIN	82.06	65.59	83.03	64.34	84.18	0.98	99.70	1.24	99.72	33.04	91.66
MDS	82.06	68.43	77.91	62.43	81.68	18.74	95.26	17.22	95.99	41.71	87.71
RMDS	82.06	48.48	87.43	43.99	88.97	9.16	96.67	6.52	97.95	27.04	92.76
EBO	82.06	59.63	82.80	56.77	84.27	4.42	98.92	5.09	98.64	31.48	91.16
MLS	82.06	57.97	83.62	55.14	85.08	4.00	98.97	4.31	98.83	30.36	91.63
KNN	82.06	77.11	73.14	70.94	78.26	0.05	99.99	0.24	99.92	37.09	87.83
ASH	82.06	85.97	66.87	84.12	69.85	3.12	99.15	12.55	96.68	46.44	83.14
SHE	82.06	87.54	62.22	85.12	65.64	4.54	98.93	7.61	97.93	46.20	81.18

REFERENCE

[1] Bitterwolf, J., Müller, M., & Hein, M. (2023). In or Out? Fixing ImageNet Out-of-Distribution Detection Evaluation. Proceedings of the 40th International Conference on chine Learning, Proceedings of Machine Learning Research, 202, 2471-2506. PMLR.

[2] Pramuditha Perera and Vishal M Patel, "Deep transfer learning for multiple class novelty detection," in Proceedings of the ieee/cvf conference on computer vision and tern recognition, 2019, pp. 11544–11552.

[3] Zhang, J., Inkawhich, N., Linderman, R., Chen, Y., & Li, H. (2023). Mixture Outlier Exposure: Towards Out-of-Distribution Detection in Fine-grained Environments. In ceedings of the 2023 IEEE Winter Conference on Applications of Computer Vision (WACV 2023). DOI: 10.1109/WACV56688.2023.00549.