1 URLs

Github: https://github.com/ServiceNow/broad
     Huggingface hub: https://huggingface.co/datasets/ServiceNow/PartialBROAD

2 Metadata

```
license: cc-by-4.0
dataset_info:
  features:
    - name: date_captured
      dtype: string
    - name: coco_url
      dtype: string
    - name: license_name
      dtype: string
    - name: license_url
      dtype: string
    - name: coco_id
      dtype: string
    - name: image
      dtype: image
    - name: label
      dtype: int64
    - name: flickr_url
      dtype: string
  splits:
    - name: clean
      num_bytes: 333801279.747
      num_examples: 36157
    - name: cocomagenet
      num_bytes: 306403223
      num_examples: 2000
    - name: cocomagenet_mono
      num_bytes: 18956338
      num_examples: 2000
    - name: synthetic_gan
      num_bytes: 242598071.454
      num_examples: 24999
    - name: synthetic_diffusion
      num_bytes: 283705025
      num_examples: 25000
    - name: adversarial_autoattack_resnet
      num_bytes: 40058245
      num_examples: 5000
    - name: adversarial_autoattack_vit
      num_bytes: 35610460
      num_examples: 5000
    - name: adversarial_pgd_resnet
      num_bytes: 65806170
      num_examples: 5000
    - name: adversarial_pgd_vit
      num_bytes: 51803590
      num_examples: 5000
  download_size: 1432934722
```

```
  dataset_size: 1378742402.201
pretty_name: BROAD
size_categories:
  - 10K<n<100K
tags:
  - imagenet
  - OOD detection
  - distribution shift
```

3 Author Statement

The authors bear all responsibility in case of violation of rights.

4 Hosting, licensing, and maintenance plan

The external splits are hosted on their original platforms and we provide on
Github the id lists and code required to obtain their data. The internal splits
are hosted via Huggingface and we will ensure it remains accessible.

5 Dataset documentation

We provide below a data card built with the data card framework (`https://
sites.research.google/datacardsplaybook/`) recommended by the NeurIPS
2024 Datasets and Benchmarks Track Call for Papers.

Benchmarking Resilience Over Anomaly Diversity (BROAD)

This dataset is a collection of images from different distribution shifts, based on the ImageNet-2012 dataset. It is designed to evaluate the capability of out-of-distribution detection methods in detecting a broad variety of distribution shifts.

DATASET LINK

https://huggingface.co/datasets/ServiceNow/PartialBROAD

DATA CARD AUTHOR(S)

Charles Guille-Escuret: Lead author

Pierre-André Noël: Contributor

Ioannis Mitliagkas: Contributor

David Vazquez: Contributor

Joao Monteiro: Supervising author

Authorship

Publishers

PUBLISHING ORGANIZATION(S)	INDUSTRY TYPE(S)	CONTACT DETAIL(S)
ServiceNow Research	Corporate - Tech	**Publishing POC:** Joao Monteiro **Affiliation:** ServiceNow **Contact:** joao.monteiro@servicenow.com

Dataset Overview

DATA SUBJECT(S)	DATASET SNAPSHOT	CONTENT DESCRIPTION
Data about places and objects Synthetically generated data		Each instance is a JPEG image. For splits with valid ImageNet classes, the instances are organized in folders, one per class.

Size of Dataset	1906 MB
Number of Instances	139789
Labeled Classes	1000

Sensitivity of Data

SENSITIVITY TYPE(S)	FIELD(S) WITH SENSITIVE DATA	SECURITY AND PRIVACY HANDLING
None	None	None

RISK TYPE(S)	SUPPLEMENTAL LINK(S)	RISK(S) AND MITIGATION(S)
No Known Risks		

Dataset Version and Maintenance

MAINTENANCE STATUS	VERSION DETAILS	MAINTENANCE PLAN
Limited Maintenance (The data will not be updated, but any technical issues will be addressed.)	**Current Version:** 1.0 **Last Updated:** 08/2023 **Release Date:** 08/2023	This dataset will be hosted on Huggingface.
	NEXT PLANNED UPDATE(S)	EXPECTED CHANGE(S)
	None	None

Motivations & Intentions

Motivations

PURPOSE(S)	DOMAIN(S) OF APPLICATION	MOTIVATING FACTOR(S)
Research	*Machine Learning, Computer Vision, OOD detection*	*OOD detection methods are typically only evaluated on unknown labels. This dataset aims to provide a more realistic evaluation for detecting unexpected image inputs.*

Intended Use

DATASET USE(S)	SUITABLE USE CASE(S)	UNSUITABLE USE CASE(S)
Safe for research use	**Measurement of the detection metrics (AUROC, FP@95 etc) of an OOD detection method.**	
	RESEARCH AND PROBLEM SPACE(S)	CITATION GUIDELINES

	It is unclear how well detection methods perform on distribution shifts other than unknown classes, that still may occur in real-world applications.	**BiBTeX:**
		
		``
		@misc{guilleescuret2023expecting,
		     title={Expecting The Unexpected: Towards Broad Out-Of-Distribution Detection},
		     author={Charles Guille-Escuret and Pierre-André Noël and Ioannis Mitliagkas and David Vazquez and Joao Monteiro},
		     year={2023},
		     eprint={2308.11480},
		     archivePrefix={arXiv},
		     primaryClass={cs.LG}
		}
		``

Access, Retention, & Wipeout

Access

ACCESS TYPE	DOCUMENTATION LINK(S)	PREREQUISITE(S)
External - Open Access	https://github.com/ServiceNow/broad https://huggingface.co/datasets/ServiceNow/PartialBROAD	*Downloading images from external splits requires respecting the conditions of each of these splits. For instance, downloading clean ImageNet images requires getting approval from the ImageNet foundation.*

Retention

	DURATION	POLICY SUMMARY
	No limitation.	

Provenance

Collection

METHOD(S) USED	METHODOLOGY DETAIL(S)	SOURCE DESCRIPTION(S)
Artificially Generated Taken from other existing datasets	**Artificially Generated:** **Source:** Generated by adversarial attacks from AutoAttack on pretrained ResNet50 and ViT-b/16, and by generation from conditional BigGan and stable diffusion. More details on https://huggingface.co/datasets/ServiceNow/PartialBROAD **Platform:** N/A **Is this source considered sensitive or high-risk?** No **Primary modality of collected data:** Image Data **Taken from other existing datasets:** **Source:** ImageNet-2012, iNaturalist, MS COCO, ImageNet-O, ImageNet-C, OpenImage-O **Is this source considered sensitive or high-risk?** No **Primary modality of collected data:** Image Data	**Artificially Generated:** - Adversarial attacks - Synthetic samples **Taken from other existing datasets:** - Subset of images from ImageNet-2012, iNaturalist, MS COCO, ImageNet-C, and full dataset from ImageNet-O, OpenImage-O
COLLECTION CADENCE	DATA INTEGRATION	DATA PROCESSING

Static (Data was collected once from single or multiple sources.)		No specific processing

Collection Criteria

DATA SELECTION	DATA INCLUSION	DATA EXCLUSION
Available datasets for corruptions, unknown labels and multiple labels Generated data for adversarial attacks and synthetic data	Clean: random subset of ImageNet validation set Adversarial attacks: all attacks for PGD, only the best attack model for AutoAttack Synthetic: all data included Unknown classes: all data included CoComageNet: please refer to :https://huggingface.co/datasets/ServiceNow/PartialBROAD	No Exclusions

Relationship to Source

USE & UTILITY(IES)	BENEFIT AND VALUE(S)	LIMITATION(S) AND TRADE-OFF(S)
Evaluation of detection scores	**Existing datasets:** agglomeration of images from different types of distribution shifts for broad evaluation **Synthetic data:** Standardization of benchmarking by providing a fixed set of synthetic data for reproducible evaluations, and representation of rarely evaluated distribution shift types.	**All splits are distribution shifts for ImageNet-2012. No other in-distribution is available.**

Extended Use

Use with Other Data

SAFETY LEVEL	KNOWN SAFE DATASET(S) OR DATA TYPE(S)	BEST PRACTICES
Safe to use with other data	**Any dataset**	Add novel splits that can be interpreted as distribution shifts from ImageNet-2012. Keep distribution shift types separated.
	KNOWN UNSAFE DATASET(S) OR DATA TYPE(S)	LIMITATION(S) AND RECOMMENDATION(S)

	No known unsafe datasets or data types.	The in-distribution is limited to ImageNet-2012

Forking & Sampling

SAFETY LEVEL	ACCEPTABLE SAMPLING METHOD(S)	BEST PRACTICE(S)
Should not be forked and/or sampled		Evaluate detection scores on all data for comparability and standardization
	RISK(S) AND MITIGATION(S)	LIMITATION(S) AND RECOMMENDATION(S)
All data should be included in the evaluation to guarantee fair evaluations		

Use in ML or AI Systems

DATASET USE(S)	NOTABLE FEATURE(S)	USAGE GUIDELINE(S)
Testing		**Compare detection score to the in-distribution (clean) and that of each of the other splits. Measure detection metrics such as AUROC or FPR@95**

Known Applications & Benchmarks

ML APPLICATION(S)	EVALUATION RESULT(S)	EVALUATION PROCESS(ES)
Out-of-distribution Detection	Please refer to arxiv.org/pdf/2308.11480	Detection score of each benchmarked detection method were computed on each split, and detection metric with respect to the clean split deduced.