Make an Omelette with Breaking Eggs:
Zero-Shot Learning for Novel Attribute Synthesis

Yu-Hsuan Li*[1], Tzu-Yin Chao*[1], Ching-Chun Huang[1], Pin-Yu Chen[2], Wei-Chen Chiu[1]

National Chiao Tung University[1]   IBM Research[2]

* The authors contributed equally to this work

Motivation

Many zero-shot learning methods for classification nowadays rely on the auxiliary information based on attributes. However, attribute annotations are costly. We thus start to think: "Given annotations for only few attributes (i.e., seen attributes), can we derive zero-shot learning for novel attribute detectors and use them to automatically annotate the dataset for labeling efficiency?"

Contribution

We show on the CUB dataset that, given only **32 attributes with manual annotations**, our ZSLA is able to synthesize 207 novel attribute detectors to provide high-quality annotations for the dataset. By utilizing the auto-annotated attributes, **generalized zero-shot classification algorithms can achieve comparable or even better performance than that using 312 manually-annotated attributes.**

Automatically Novel Attribute Annotations For A Dataset Using ZSLA

1. Training the seen attribute detectors.
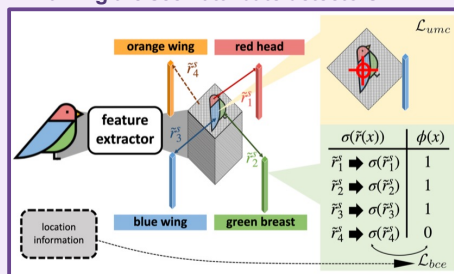
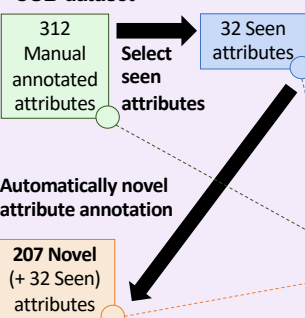2. Training the intersection and union network.
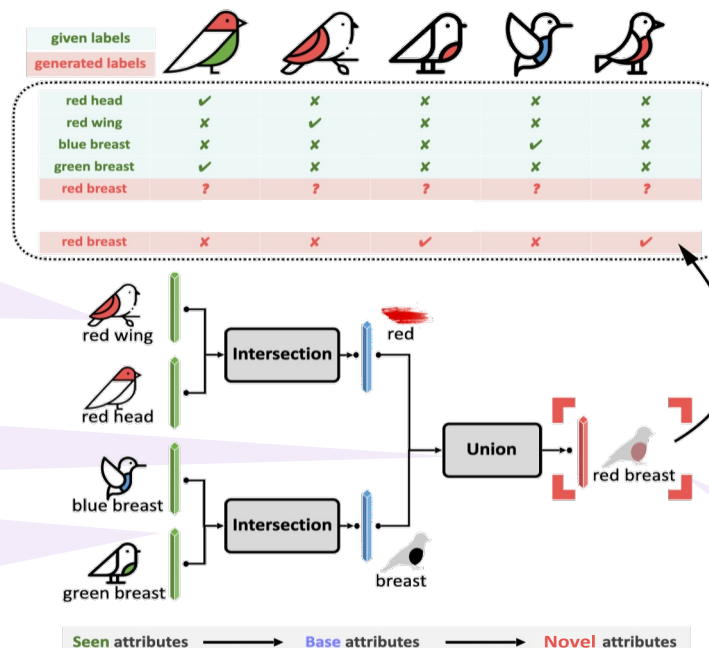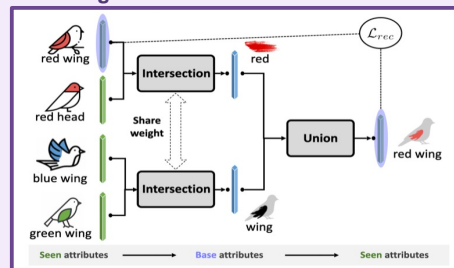
0. Annotating a few attributes manually at a start during constructing a dataset.

4. Automatically providing novel attribute annotations for the dataset.

3. Synthesize the novel attribute detectors.

Seen attributes → Base attributes → Novel attributes

Experimental Results

+ Attribute re-annotation on CUB dataset

1. Evaluation of synthesized novel attribute detectors on CUB dataset for **attribute classification (mAUROC)**, **retrieval (mAP@50)**, and **localization (mLA)**.

	mAUROC			mAP@50			mLA		
# Seen Attributes	32	64	96	32	64	96	32	64	96
A-LAGO	0.600	0.612	0.627	0.173	0.180	0.222	0.782	0.787	0.795
A-ESZSL	0.626	0.614	0.632	0.223	0.200	0.234	0.756	0.769	0.756
Our ZSLA	**0.689**	**0.704**	**0.717**	**0.320**	**0.327**	**0.329**	**0.846**	**0.860**	**0.867**

2. Re-annotation experiment on CUB dataset: Comparing **GZSL performance of 4 different algorithms** trained by utilizing either fully manual annotated attributes or the ones **re-annotated+** by different methods.

	CADAVAE			TFVAEGAN			ALE			ESZSL		
	S	U	H	S	U	H	S	U	H	S	U	H
Manual (# Seen Attributes=32)	42.9	27.3	33.4	45.5	31.2	37.1	26.4	9.2	13.7	29.8	10.8	15.9
Manual (# Seen Attributes=312)	53.5	51.6	52.4	64.7	52.8	58.1	62.8	23.7	34.4	63.8	12.6	21.0
A-LAGO	45.4	55.4	49.9	57.4	53.0	55.1	51.8	27.2	35.6	49.7	17.1	25.4
A-ESZSL	41.5	48.7	44.8	56.0	48.5	52.0	49.7	17.1	25.4	61.3	9.2	16.0
Our ZSLA (# Seen Attributes=32, # Novel Attributes=207)	50.3	56.4	53.2	59.0	55.9	57.4	52.4	27.5	36.1	65.1	16.4	26.2

3. Examples of **novel attribute retrieval and localization**. Each set shows the top-5 retrieved images and their response maps for a novel attribute.