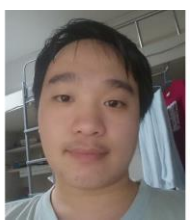
ICME 2022, Taipei, Taiwan

Generative and Adaptive Multi-Label Generalized Zero-Shot Learning

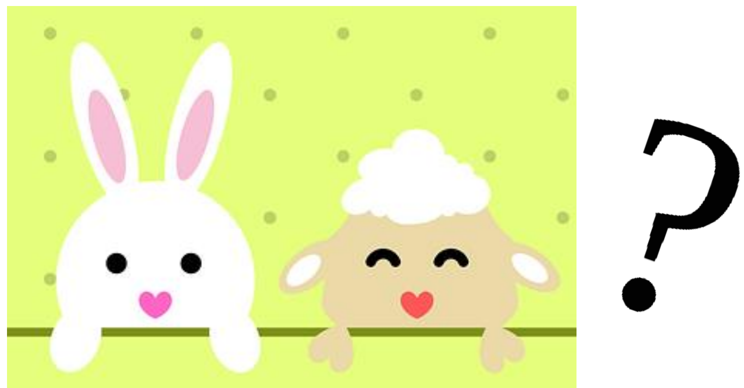
Kuan-Ying Chen and Mei-Chen Yeh



Department of Computer Science and Information Engineering
National Taiwan Normal University

Multi-label classification

Goal: Recognize one or multiple objects in one image

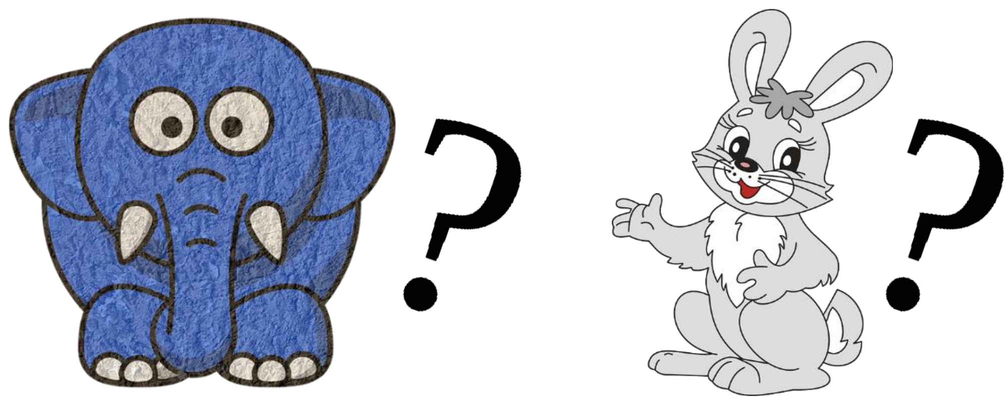


Seen classes

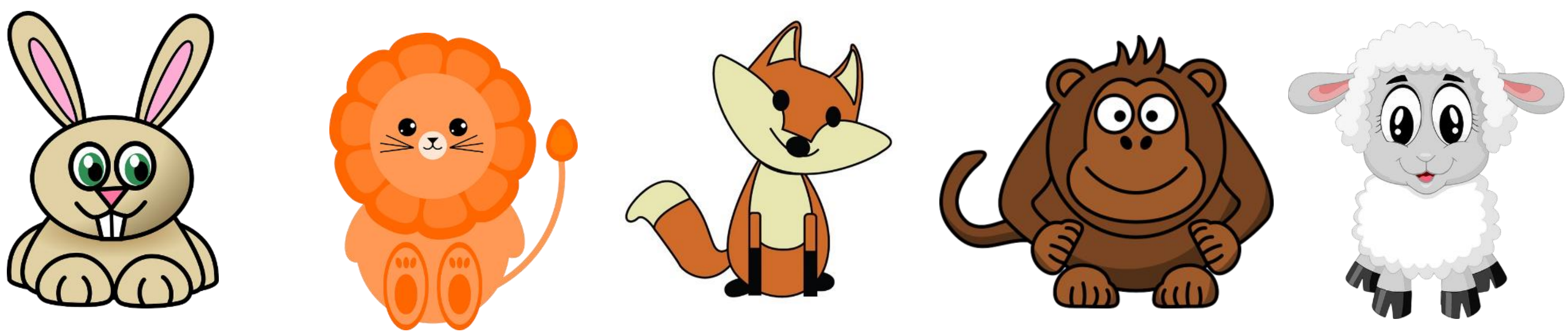
Images are from https://pixabay.com/.

Generalized zero-shot learning

Goal: Recognize objects whose instances may not have been seen during training

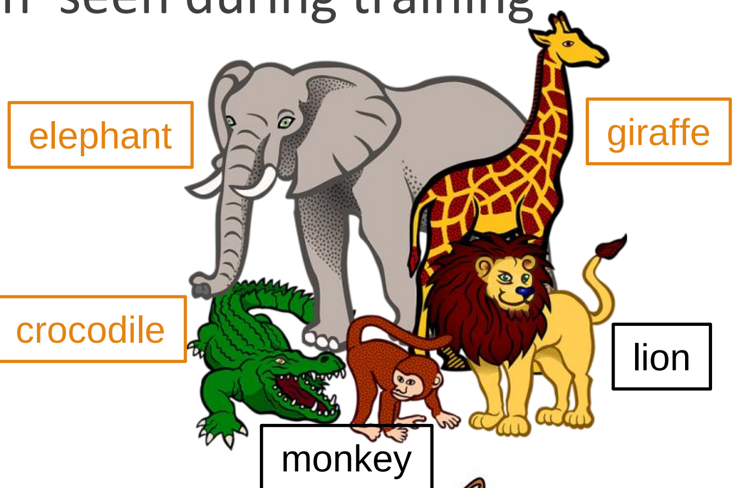


Seen classes

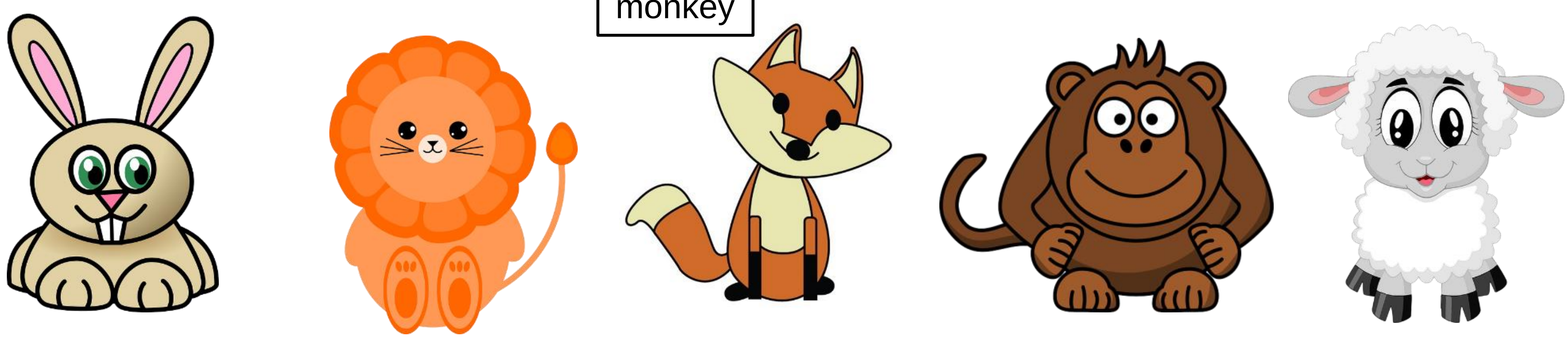


Multi-label generalized zero-shot learning

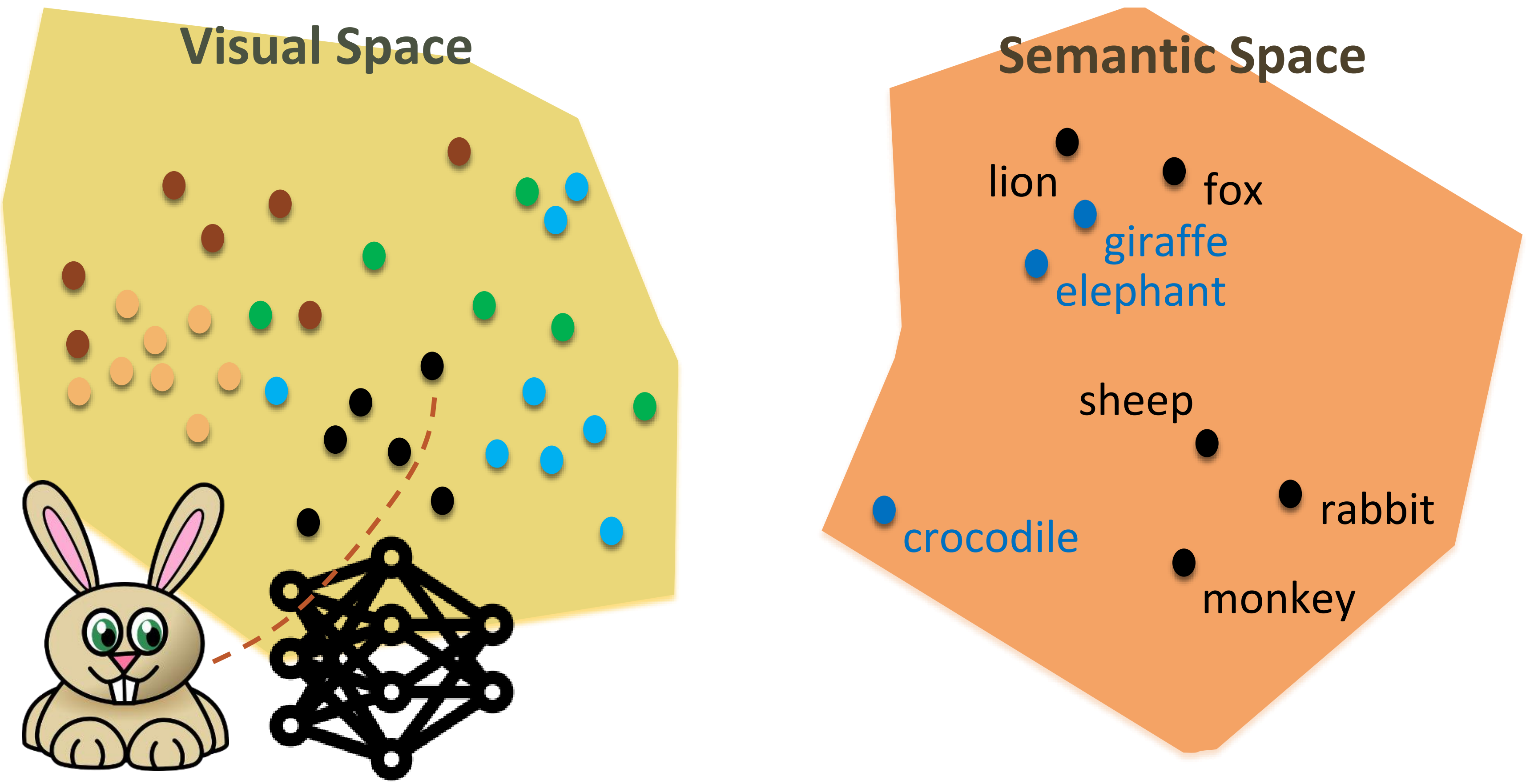
Goal: Recognize one or multiple objects whose instances may not have been seen during training



Seen classes



Visual and semantic embeddings



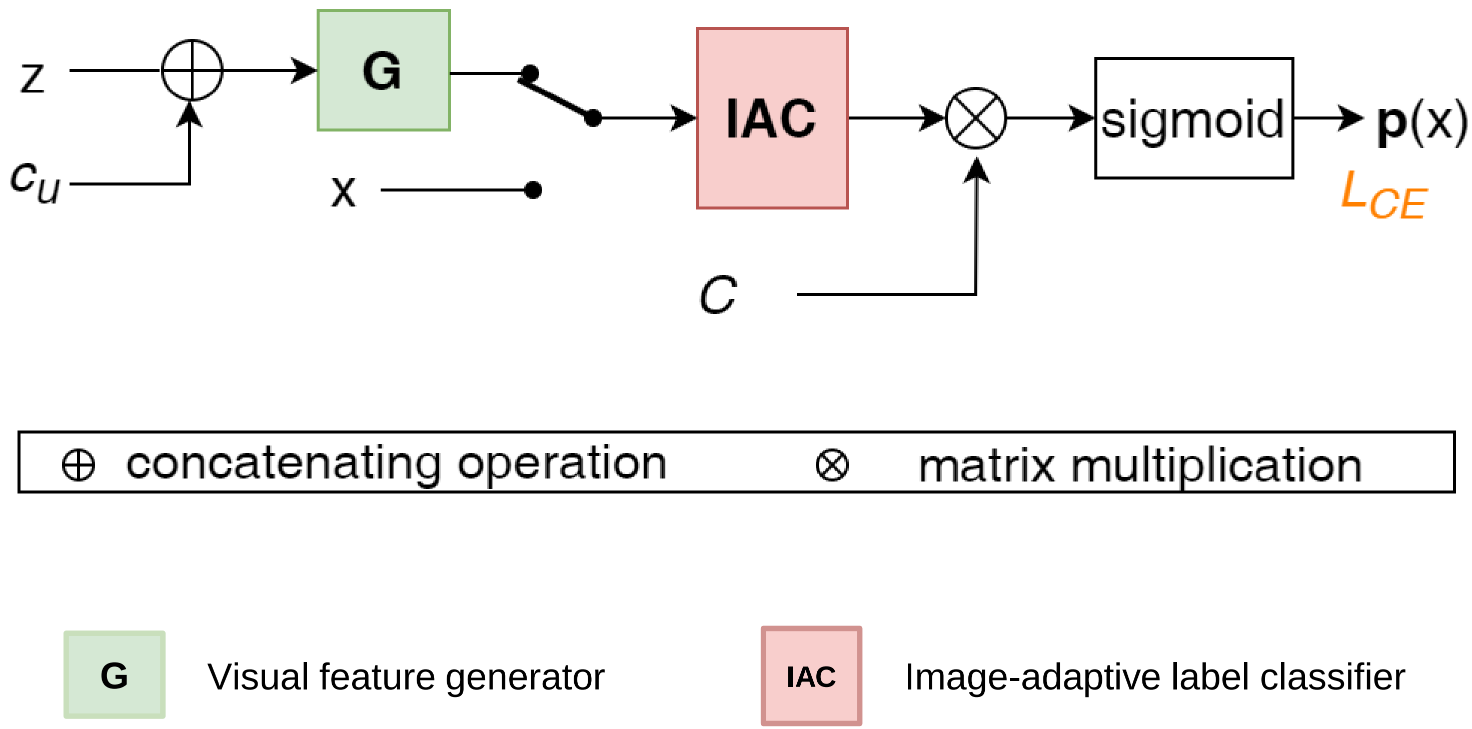
Generative methods for GZSL

- Synthesize training samples for unseen classes
- Real (seen) + synthesized (unseen) visual features => fully-observed training set for both seen and unseen classes
- Successful for single-label GZSL
- Not trivial for multi-label setting because the location of each label in one image is now known
- *How to synthesize **multi-label** visual features from **multi-label** images?*

Contributions

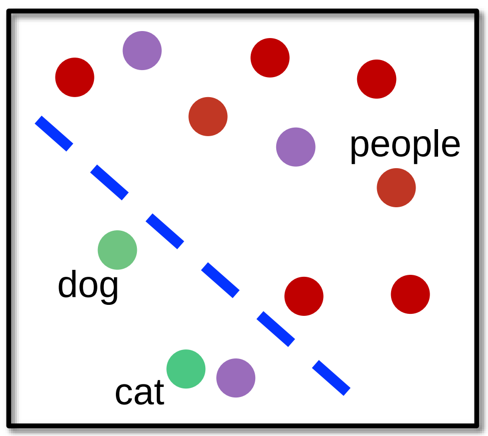
- We present a new approach based on the generative paradigm for multi-label GZSL.
- We apply the concept of converting an image into a label classifier. The adaptive nature of the method facilitates the integration of a single-label feature generating model for creating multi-label features from multi-label images.

Approach


z
c_u
G
x
IAC
C
sigmoid
p(x)
L_{CE}
⊕ concatenating operation
⊗ matrix multiplication
G
Visual feature generator
IAC
Image-adaptive label classifier

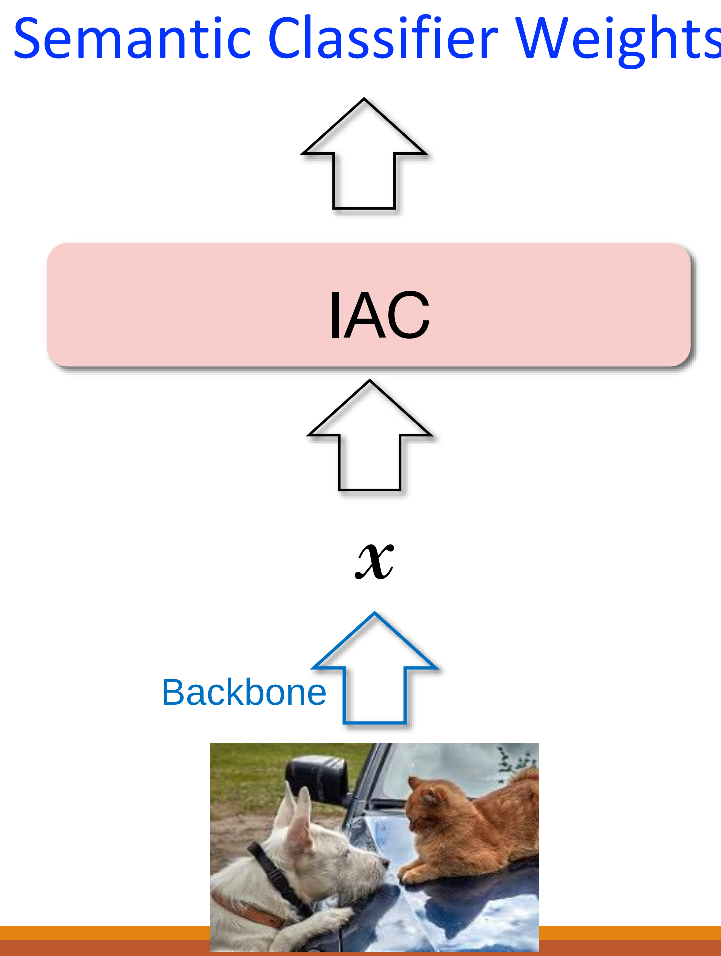
IAC

Image-adaptive label classifier



Semantic Space W

- positive seen labels
- negative seen labels
- unseen labels



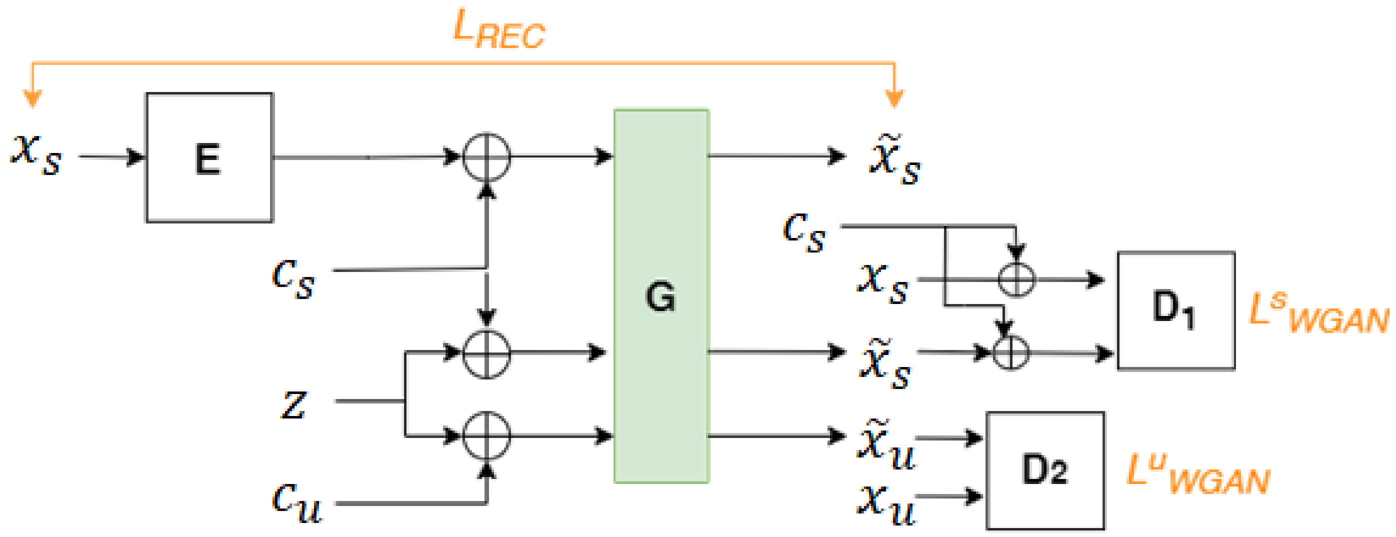
IAC

Highlights

- Flexible: images of the same class can have different semantic classifiers!
- Facilitate the task of feature generation
  - Decompose a multi-label training sample into multiple image-label pairs
  - Process a label at a time

G

Visual feature generator



Yongqin Xian, Sauabh Sharma, Bernt Schiele, and Zeynep Akata, “F-VAEGAN-D2: A feature generating framework for any-shot learning,” in *IEEE CVPR*, 2019.

Experimental results

PASCAL VOC

Method	Task	MiAP	mi-F1	ma-F1
CONSE [22]	ZSL	49.98	30.8	27.57
	GZSL	64.10	42.11	32.29
LabelEM [14]	ZSL	52.45	35.32	36.69
	GZSL	66.46	43.11	32.37
DMP [23]	ZSL	53.52	36.70	40.44
	GZSL	67.79	43.97	34.13
Fast0Tag [6]	ZSL	52.39	35.01	36.76
	GZSL	67.34	43.54	33.31
TAEP-C [18]	ZSL	59.22	39.84	43.77
	GZSL	69.87	44.75	35.62
Our Approach	ZSL	**62.83**	**44.63**	**44.11**
	GZSL	**70.19**	**50.46**	**52.58**

NUS-WIDE

Method	Task	K = 3			K = 5			mAP
		P	R	F1	P	R	F1	
CONSE [22]	ZSL	17.5	28.0	21.6	13.9	37.0	20.2	9.4
	GZSL	11.5	5.1	7.0	9.6	7.1	8.1	2.1
LabelEM [14]	ZSL	15.6	25.0	19.2	13.4	35.7	19.5	7.1
	GZSL	15.5	6.8	9.5	13.4	9.8	11.3	2.2
Fast0Tag [6]	ZSL	22.6	36.2	27.8	18.2	48.4	26.4	15.1
	GZSL	18.8	8.3	11.5	15.9	11.7	13.5	3.7
One Attention per Label [24]	ZSL	20.9	33.5	25.8	16.2	43.2	23.6	10.4
	GZSL	17.9	7.9	10.9	15.6	11.5	13.2	3.7
LESA (M=10) [20]	ZSL	25.7	**41.1**	31.6	19.7	**52.5**	28.7	19.4
	GZSL	23.6	10.4	14.4	19.8	14.6	16.8	5.6
Our Approach	ZSL	**26.0**	**41.1**	**31.9**	**19.9**	52.3	**28.8**	**26.3**
	GZSL	**30.2**	**13.1**	**18.3**	**25.2**	**18.3**	**21.2**	**11**

Summary

- We propose image-adaptive classification (IAC) to address the multi-label GZSL problem.
- IAC can adaptively emphasize the most discriminating dimension in semantic features to deal with intra-class visual discrepancies.
- IAC also facilities the multi-label feature generating task by a simple decomposition approach.
- The proposed method improves the state-of-the-arts on two benchmark datasets.

Questions?

More information:

http://www.csie.ntnu.edu.tw/~myeh/