IEEE INTERNATIONAL CONFERENCE ON
IMAGE PROCESSING (ICIP 2021)

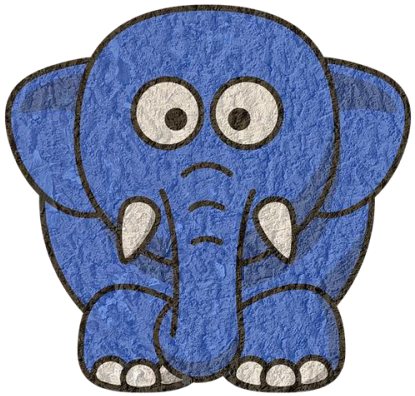
Generalized Zero-Shot Recognition through Image-Guided Semantic Classification

Fang Li and **Mei-Chen Yeh**

Dept. of Computer Science and Information Engineering
National Taiwan Normal University

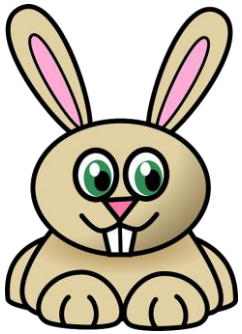
(Generalized) Zero-Shot Learning

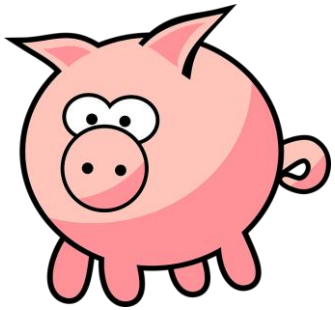
- Goal: Recognize objects whose instances may not have been seen during training



?

Seen classes





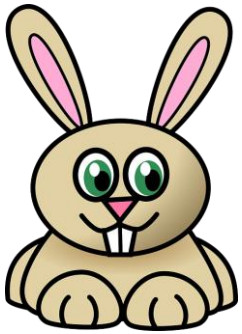
Images are from https://pixabay.com/.

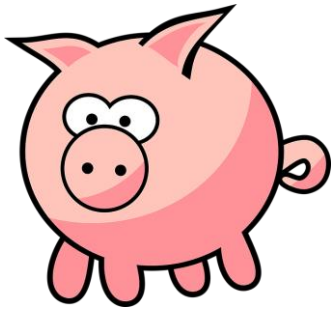
(Generalized) Zero-Shot Learning

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

?

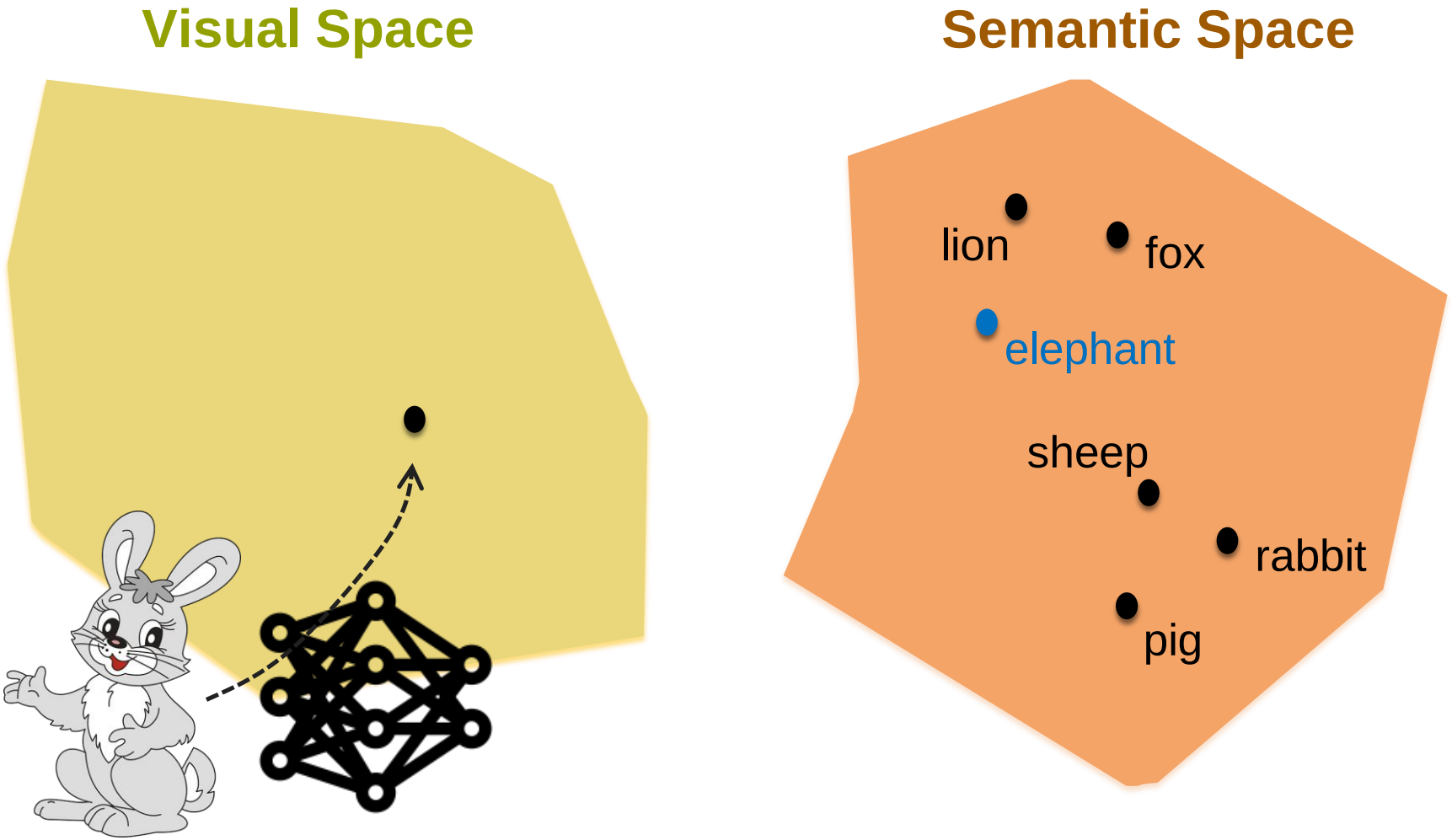
Seen classes



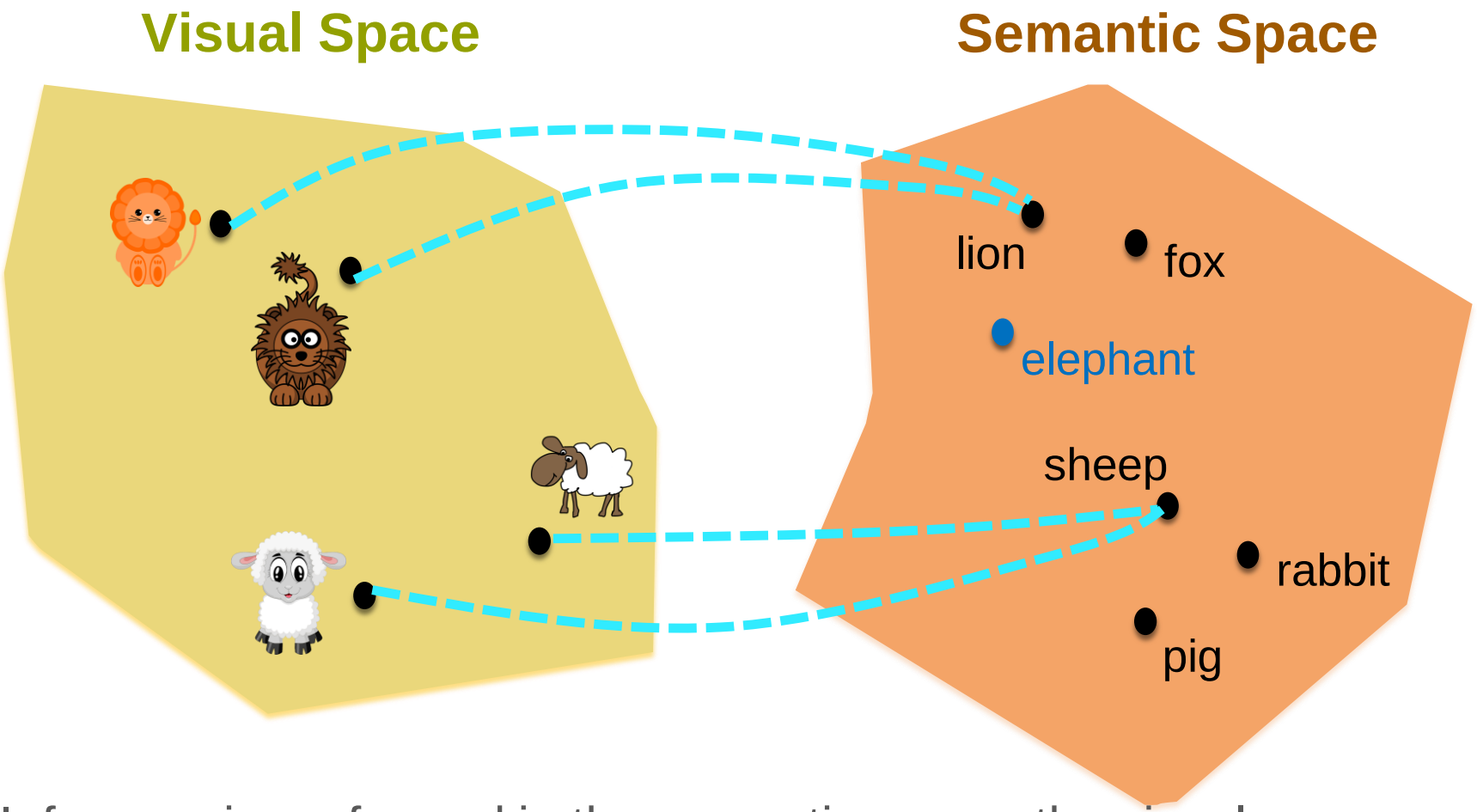


Images are from https://pixabay.com/.

Visual and Semantic Embeddings

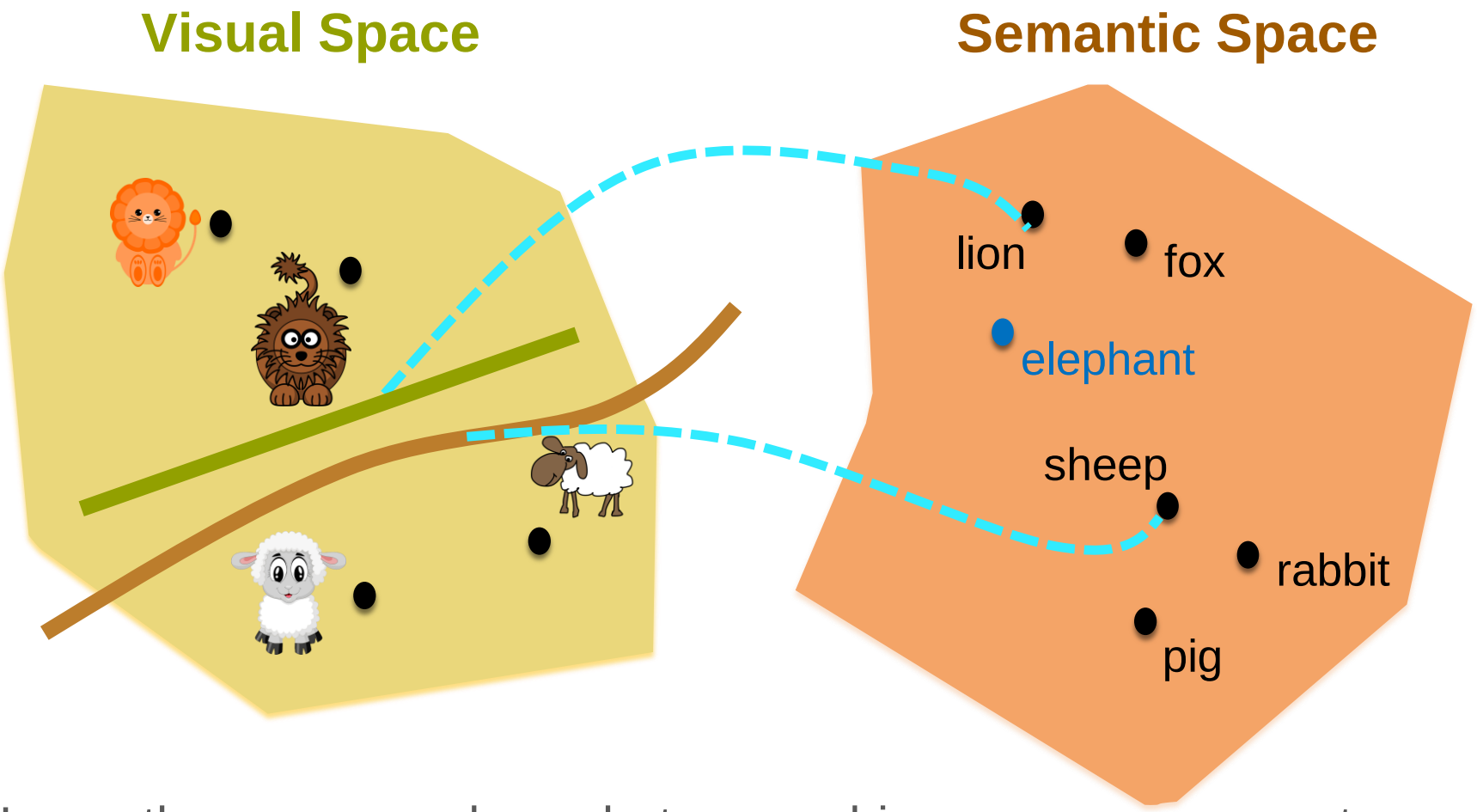


Existing Embedding-Based Methods



Inference is performed in the semantic space, the visual space, or a common space.

Existing Embedding-Based Methods

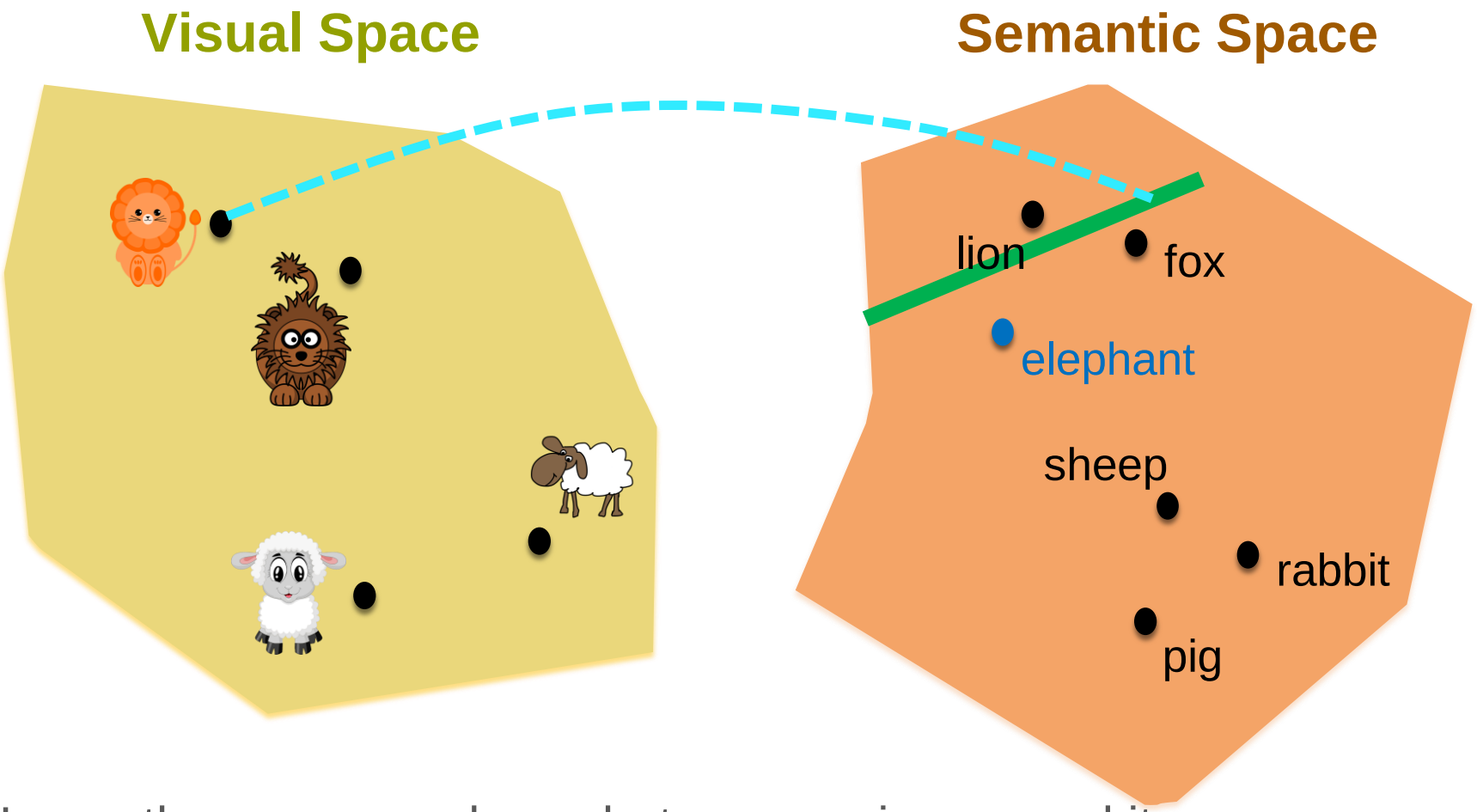


Learn the correspondence between a binary one-versus-rest image classifier and its class prototype in the semantic space

Drawbacks

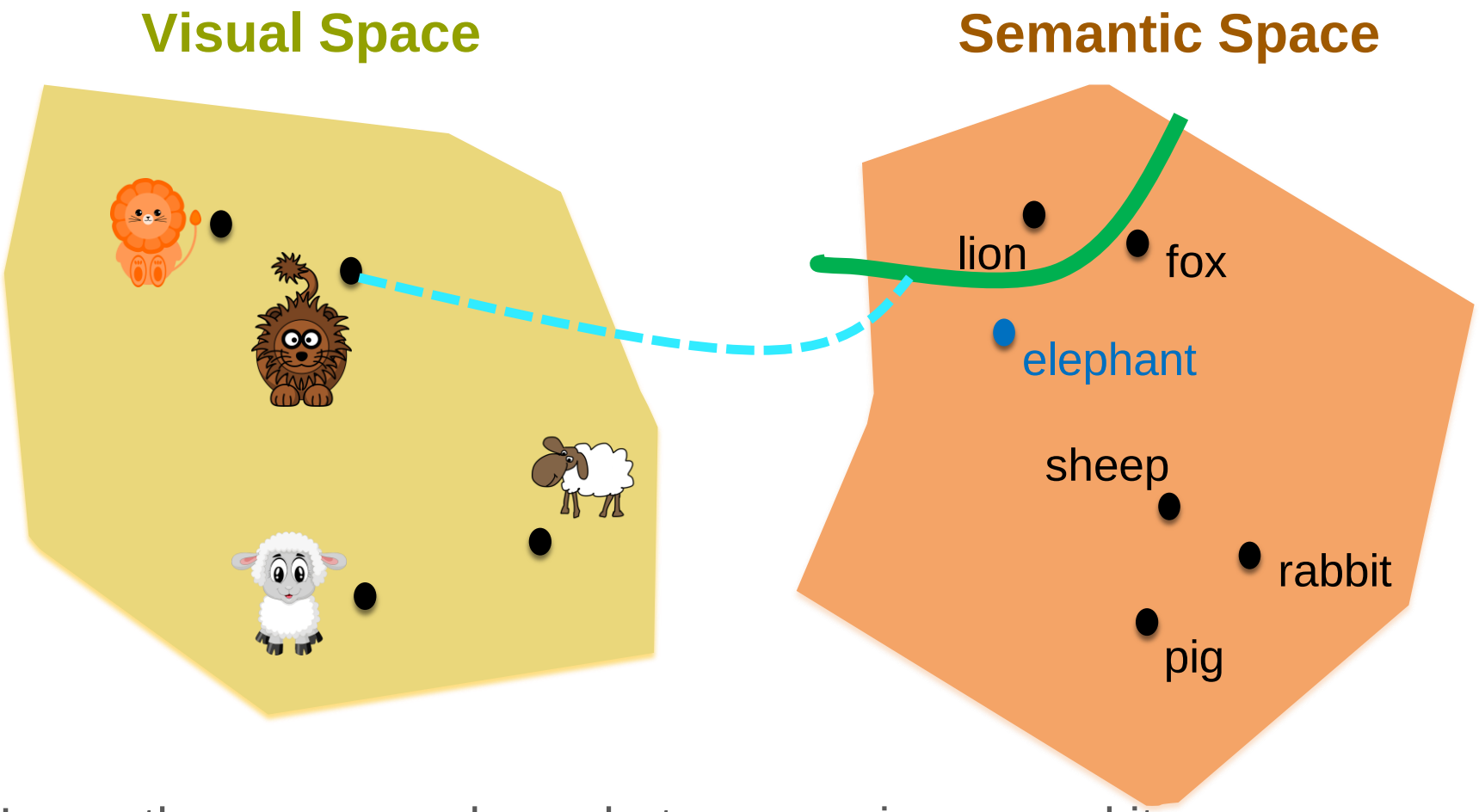
- Using a single image classifier for each class is restrictive because the manner for separating classes in both visual and semantic spaces would not be unique.
- The scale of training data for learning the correspondence is constrained to be the number of class labels.
- Each class is represented by only a single class prototype to determine where images of that class collapse inevitably.

The Proposed Method



Learn the correspondence between an image and its corresponding label classifier!

The Proposed Method



Learn the correspondence between an image and its corresponding label classifier!

Highlights of the Proposed Method

- Learns a semantic classifier from one image
- Adaptive: Label classification is conducted by an image-conditioned semantic classifier whose weights are generated on the fly.
- Has very few hyperparameters
- Can be trained end-to-end
- Compact yet powerful
- Alleviates the hubness problem

Experimental Results on Benchmark Datasets

Method	SUN			CUB			AWA2			aPY		
	acc_u	acc_s	H	acc_u	acc_s	H	acc_u	acc_s	H	acc_u	acc_s	H
LATEM [9]	14.7	28.8	19.5	15.2	57.3	24.0	11.5	77.3	20.0	1.3	71.4	2.6
DEVISE [4]	16.9	27.4	20.9	23.8	53.0	32.8	17.1	74.7	27.8	3.5	**78.4**	6.7
ESZSL [8]	11.0	27.9	15.8	14.7	56.5	23.3	5.9	77.8	11.0	2.4	70.1	4.6
SYNC [20]	7.9	**43.3**	13.4	11.5	70.9	19.8	9.7	89.7	17.5	7.4	66.3	13.3
SP-AEN [18]	24.9	38.6	30.3	34.7	70.6	46.6	23.3	90.9	37.1	13.7	63.4	22.6
PSR [13]	20.8	37.2	26.7	24.6	54.3	33.9	20.7	73.8	32.3	13.5	51.4	21.4
DCN [6]	25.5	37.0	30.2	28.4	60.7	38.7	–	–	–	14.2	75.0	23.9
AREN [21]	19.0	38.8	25.5	38.9	**78.7**	**52.1**	5.6	**92.9**	26.7	9.2	76.9	16.4
DAZLE [22]	21.7	31.9	25.8	**42.0**	65.3	51.1	**25.7**	82.5	39.2	–	–	–
IGSC	**39.4**	31.3	**34.9**	40.8	60.2	48.7	**25.7**	83.6	**39.3**	**23.1**	58.9	**33.2**

Our approach outperformed state-of-the-art embedding-based methods on most benchmark datasets!

IEEE INTERNATIONAL CONFERENCE ON
IMAGE PROCESSING (ICIP 2021)

Generalized Zero-Shot Recognition through Image-Guided Semantic Classification

Fang Li and Mei-Chen Yeh

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