

Zero-Shot Learning

Image Classification

19/11/2025



Zero-Shot Learning

CVL – Nottingham University
19th November 2026

Juan José Herrera Aranda
juanjoa@ugr.es



Funded by Knowledge Generation Projects: PID2023-149128NB-I00

Outline

- Intuition behind ZSL
- Data Organization
- Semantic Spaces
- Methods
 - Generative Methods
- Model's Performance



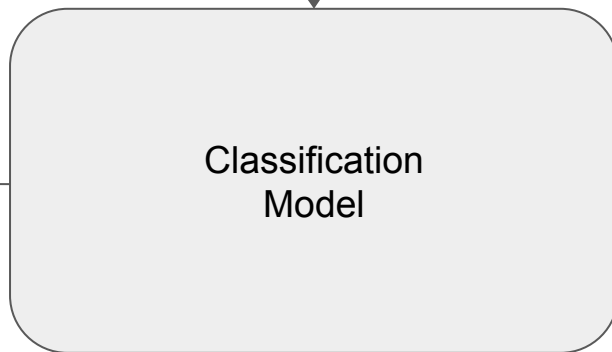
Training



Classification
Model



Training

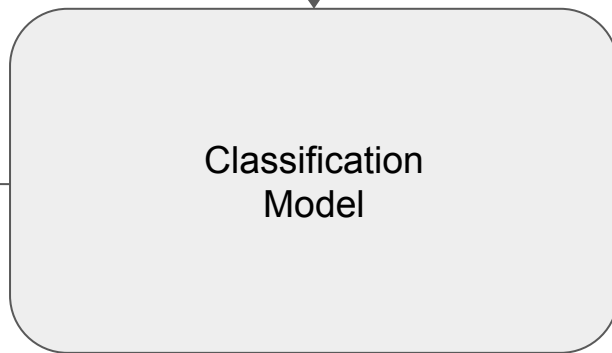


Prediction





Training

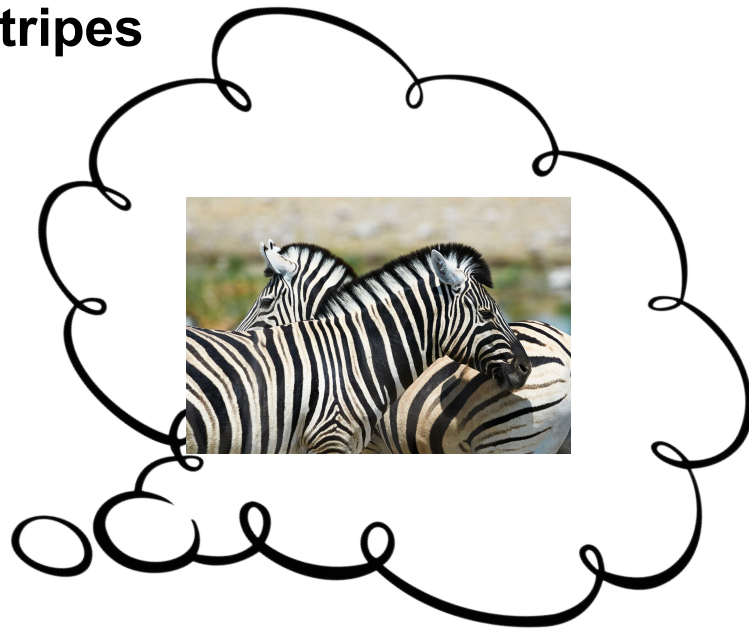


Prediction





A Zebra class is a white horse with black stripes like a tiger.



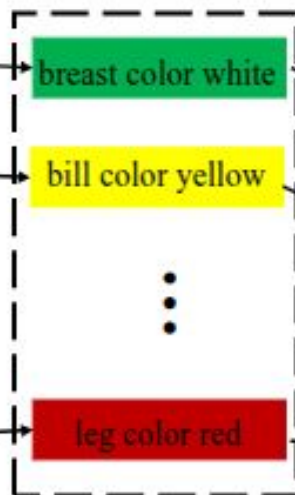
Seen Classes



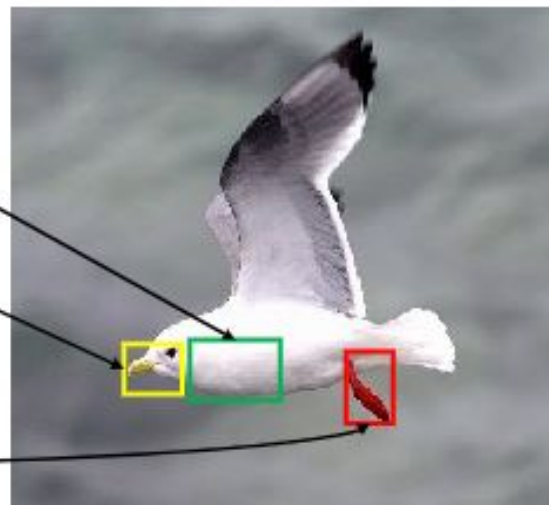
⋮



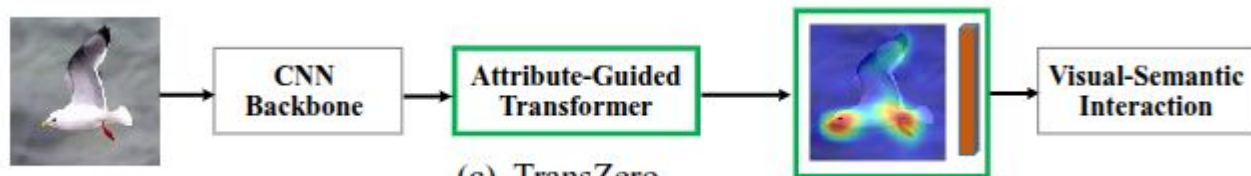
Semantic Knowledge



Unseen Class

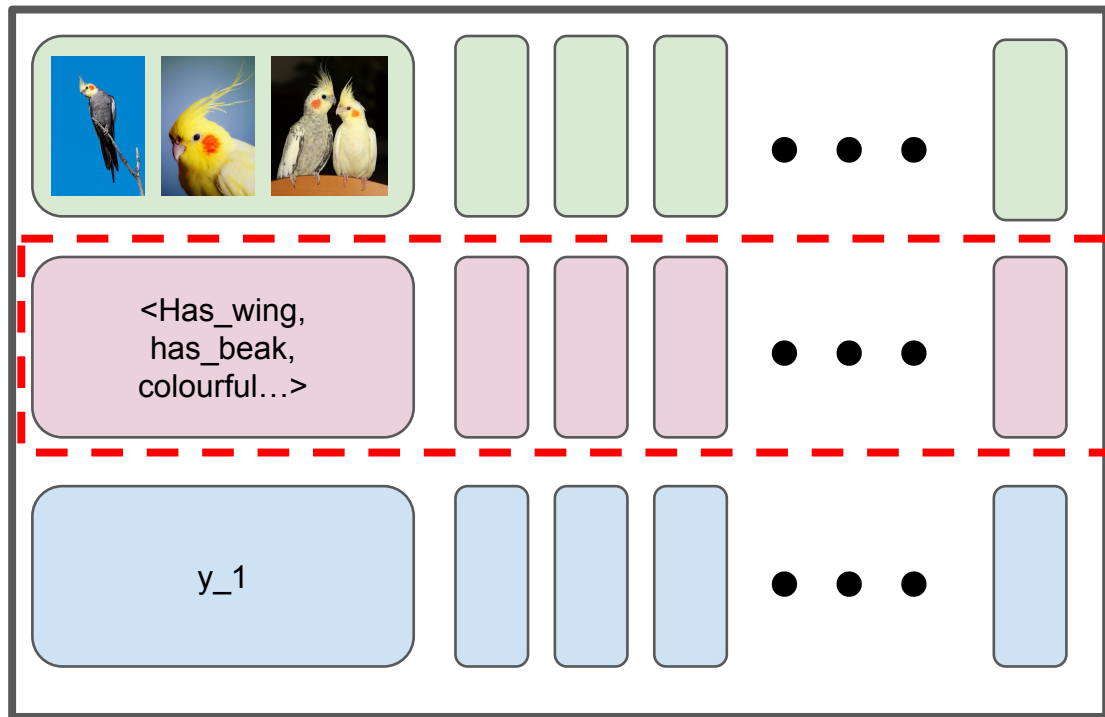


Semantic Knowledge Transfer in ZSL

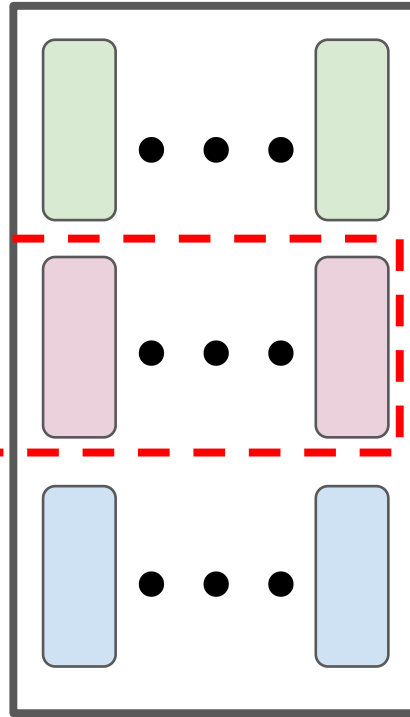


(c) TransZero

Seen Classes

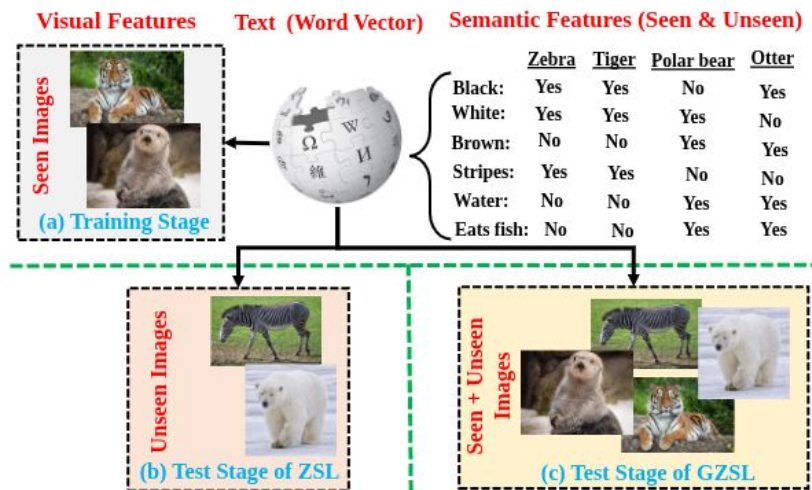


Unseen Classes



Semantic Space: The key element in ZSL

- It transfers knowledge from seen classes to unseen ones
- It should...
 - Be discriminative among classes
 - Contain transferable information



Semantic Space: Types

- Types
 - **Human-annotated**
 - Learning-based

polar bear

black: no
white : yes
brown: yes
stripes: no
water: yes
eats fish:yes



zebra

black: yes
white : yes
brown: no
stripes: yes
water: no
eats fish: no



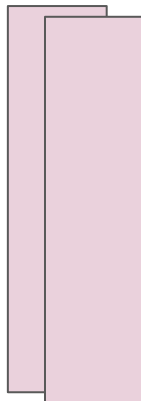
Semantic Space – Types

- Types
 - Human-annotated
 - **Learning-based**

Horses: are ungulate mammals. A horse's hearing is good, it has large ear and can rotate

Zebras: are white animals with black stripes, they have larger, rounder ears than horses

Glove,
Word2Vec,
CLIP...



polar bear

black: no
white : yes
brown: yes
stripes: no
water: yes
eats fish: yes



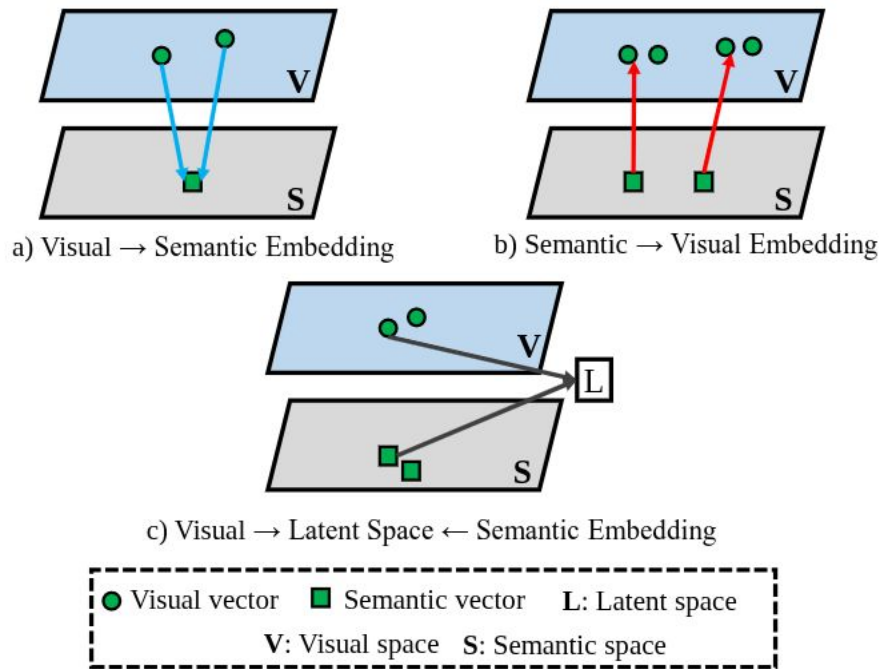
zebra

black: yes
white : yes
brown: no
stripes: yes
water: no
eats fish: no



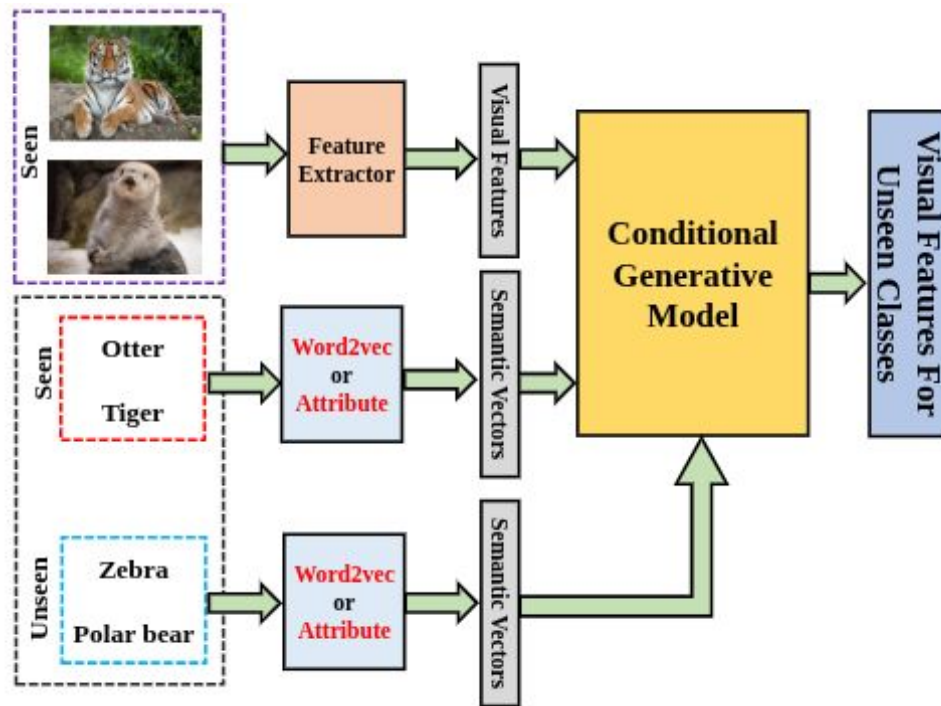
Methods

- **Embedding-based:** Project one space into another
- **Generative-based:** Generate synthetic features for unseen classes and conduct a supervised classification

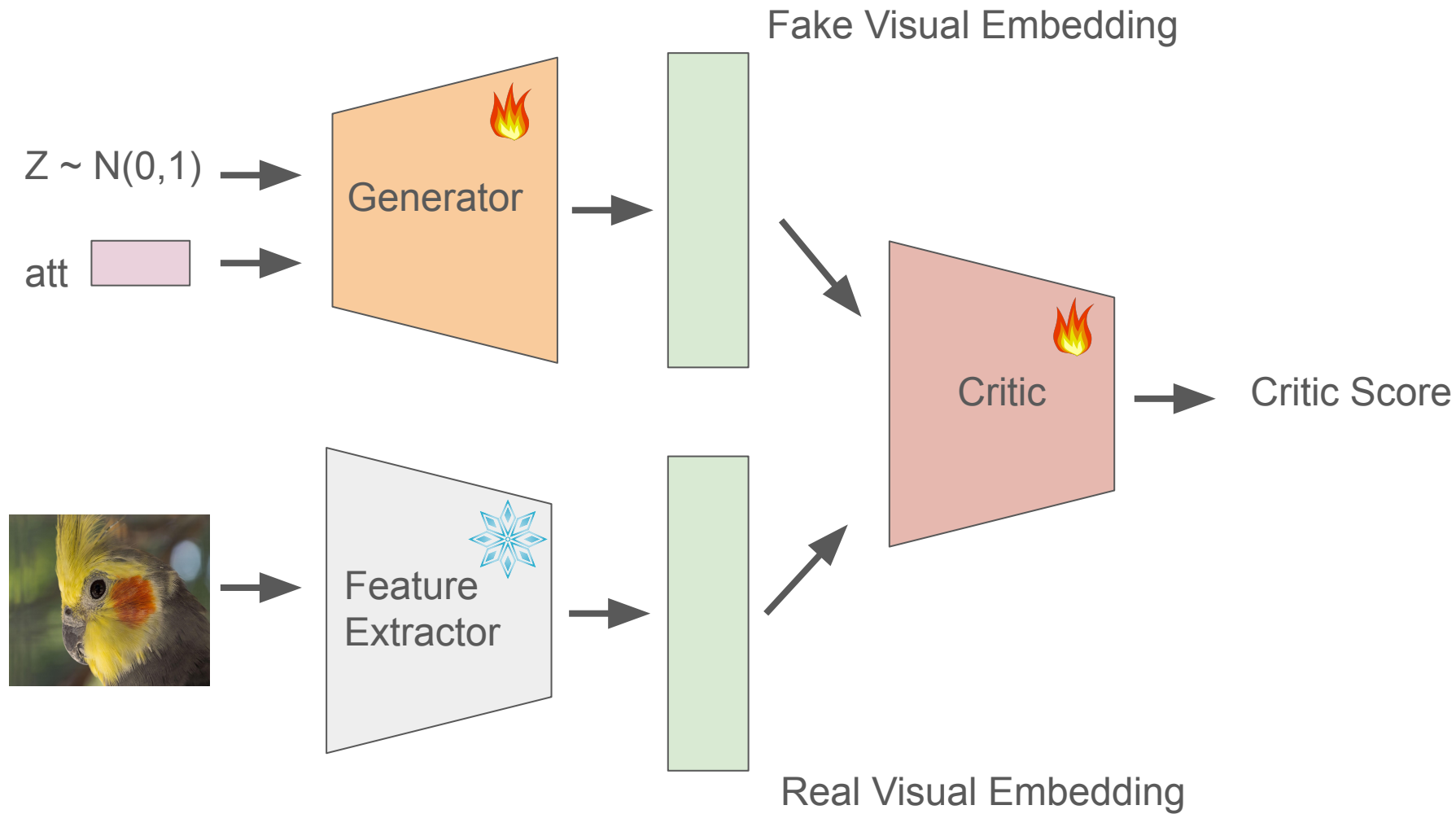


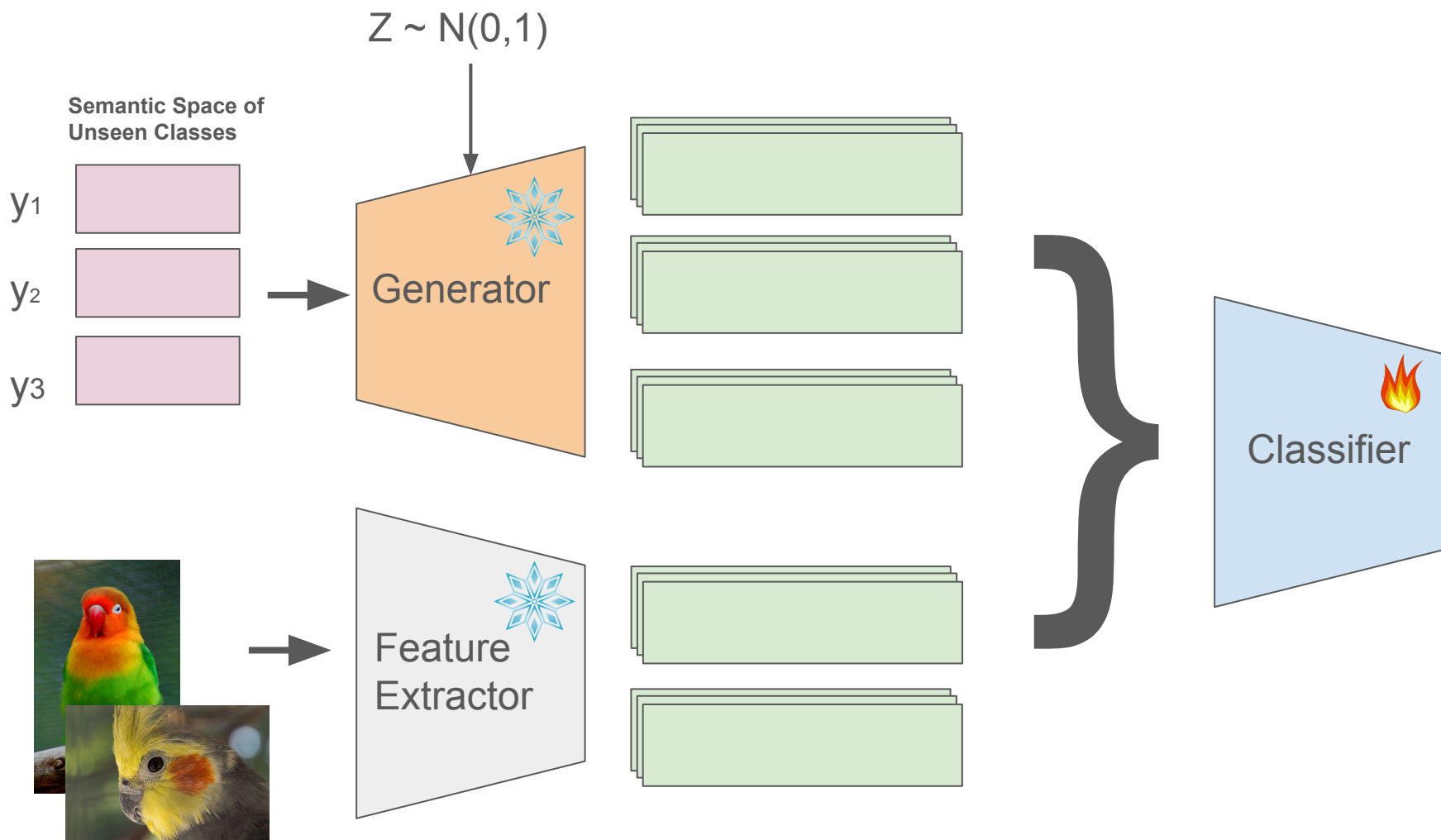
Methods

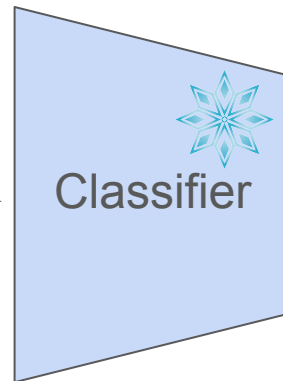
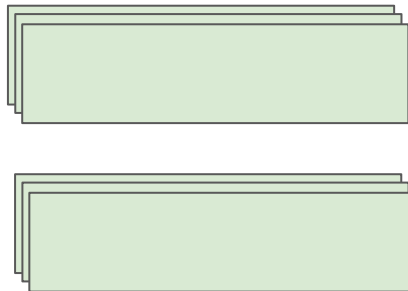
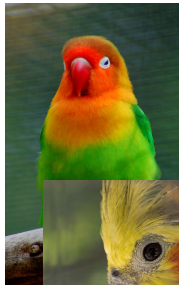
- **Embedding-based:** Project one space into another and classify using KNN
- **Generative-based:** Generate synthetic features for unseen classes and conduct a supervised classification



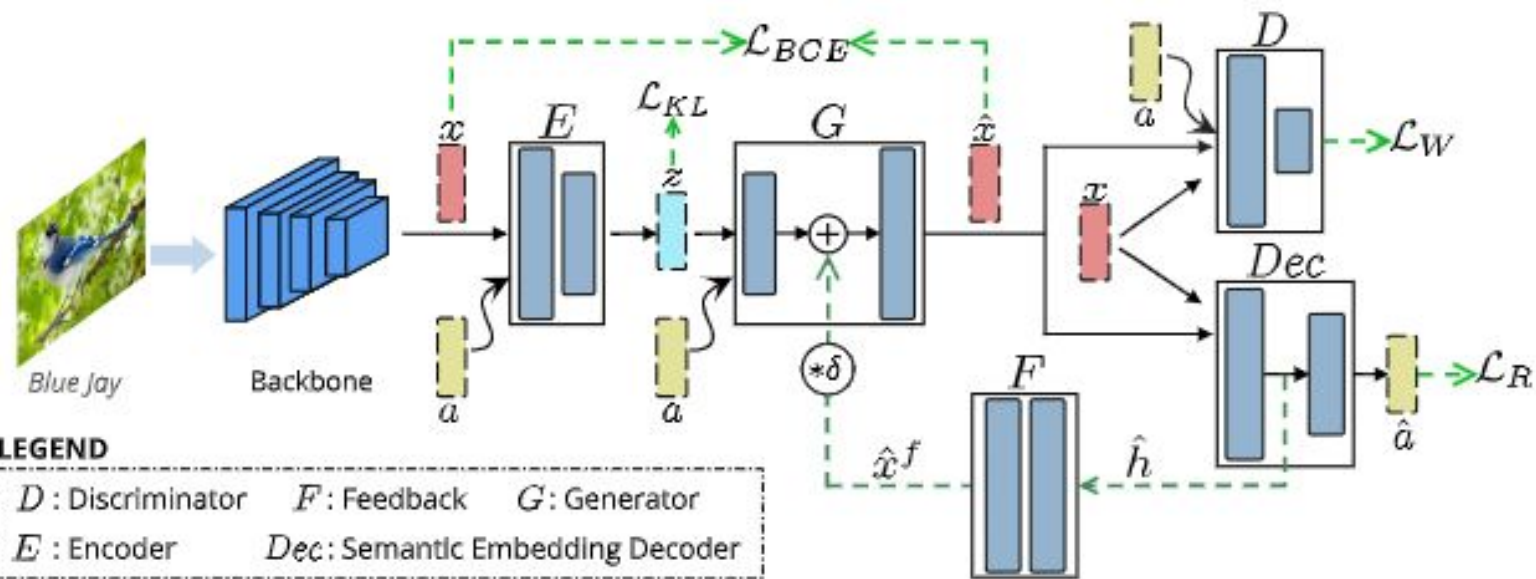
Generative-based ZSL method







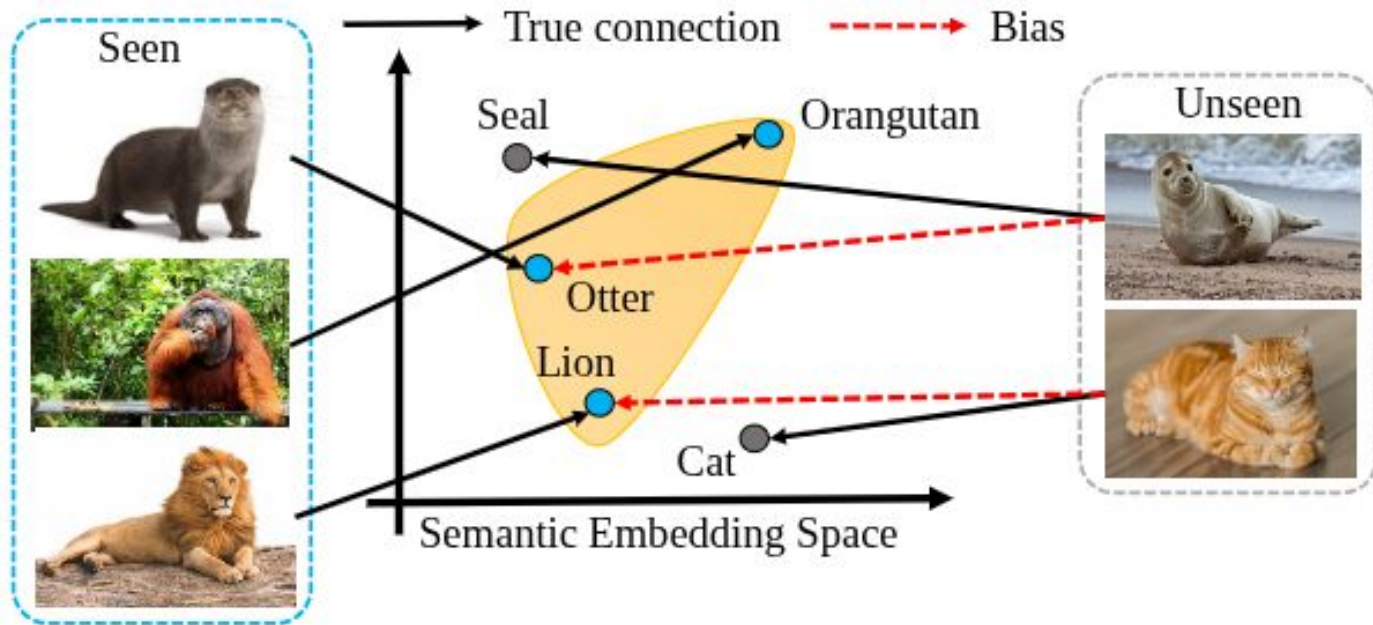
Prediction



State-of-the-art

Type	Methods	Venue	Backbone	AWA2				SUN				CUB			
				CZSL	GZSL			CZSL	GZSL			CZSL	GZSL		
				Acc	U	S	H	Acc	U	S	H	Acc	U	S	H
Embedding	PREN [52]	CVPR'19	ResNet-101	74.1	32.4	88.6	47.4	62.9	35.4	27.2	30.8	66.4	35.2	55.8	43.1
	DAZLE [22]	CVPR'20	ResNet-101	67.9	60.3	75.7	67.1	59.4	52.3	24.3	33.2	66.0	56.7	59.6	58.1
	DVBE [28]	CVPR'20	ResNet-101	–	63.6	70.8	67.0	–	45.0	37.2	40.7	–	53.2	60.2	56.5
	CN [38]	ICLR'21	ResNet-101	–	60.2	77.1	67.6	–	44.7	41.6	43.1	–	49.9	50.7	50.3
	GEM-ZSL [27]	CVPR'21	ResNet-101	67.3	64.8	77.5	70.6	62.8	38.1	35.7	36.9	77.8	64.8	77.1	70.4
	ViT-ZSL [3]	IMVIP'21	ViT-Large	–	51.9	90.0	65.8	–	44.5	55.3	49.3	–	67.3	75.2	71.0
	IEAM-ZSL [2]	DGAM'21	ViT-Large	–	53.7	89.9	67.2	–	48.2	54.7	51.3	–	68.6	73.8	71.1
	DUET [14]	AAAI'23	ViT-Base	69.9	63.7	84.7	72.7	64.4	45.7	45.8	45.8	72.3	62.9	72.8	67.5
	PSVMA [26]	CVPR'23	ViT-Base	–	73.6	77.3	75.4	–	61.7	45.3	52.3	–	70.1	77.8	73.8
Generative	f-VAEGAN-D2 [47]	CVPR'19	ResNet-101	71.1	57.6	70.6	63.5	64.7	45.1	38.0	41.3	61.0	48.4	60.1	53.6
	TF-VAEGAN [31]	ECCV'20	ResNet-101	72.2	59.8	75.1	66.6	66.0	45.6	40.7	43.0	64.9	52.8	64.7	58.1
	FREE [7]	ICCV'21	ResNet-101	–	60.4	75.4	67.1	–	47.4	37.2	41.7	–	55.7	59.9	57.7
	HSVA [8]	NeurIPS'21	ResNet-101	–	56.7	79.8	66.3	–	48.6	39.0	43.3	–	52.7	58.3	55.3
	CE-GZSL [20]	CVPR'21	ResNet-101	70.4	63.1	78.6	70.0	63.3	48.8	38.6	43.1	77.5	63.9	66.8	65.3
	FREE+ESZSL [5]	ICLR'22	ResNet-101	–	51.3	78.0	61.8	–	48.2	36.5	41.5	–	51.6	60.4	55.7
	ICCE [24]	CVPR'22	ResNet-101	72.7	65.3	82.3	72.8	–	–	–	–	78.4	67.3	65.5	66.4
	TDCSS [16]	CVPR'22	ResNet-101	–	59.2	74.9	66.1	–	–	–	–	–	44.2	62.8	51.9
	CDL + OSCO [4]	TPAMI'23	ResNet-101	–	48.0	71.0	57.1	–	32.0	65.0	42.9	–	29.0	69.0	40.6
	CLSWGAN+DSP [12]	ICML'23	ResNet-101	–	60.0	86.0	70.7	–	48.3	43.0	45.5	–	51.4	63.8	56.9
	TFVAEGAN+SHIP [43]	ICCV'23	ViT-Base	–	61.2	95.9	74.7	–	–	–	–	–	22.5	82.2	35.3
	VADS		ViT-Base	82.5	75.4	83.6	79.3	76.3	64.6	49.0	55.7	86.8	74.1	74.6	74.3

Bias towards seen classes in Generalized ZSL setting



Thank you so much :)