GenZSL: Generative Zero-Shot Learning Via Inductive Variational Autoencoder

Shiming Chen[1], Dingjie Fu[2], Salman Khan[1,3], Fahad Shahbaz Khan[1,4]

[1]Mohamed bin Zayed University of AI

[2]Huazhong University of Science and Technology

[3]Australian National University   [4]Linköping University

Contents

- Zero-Shot Learning

- Motivation

- The Proposed GenZSL

- Advantages
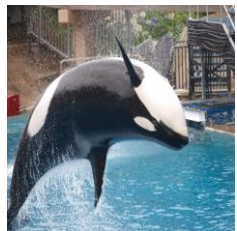
Zero-Shot Learning

What is zero-shot learning?

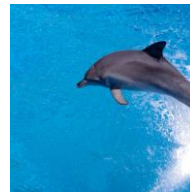Training Time

Test Time

tiger

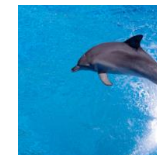zebra

killer whale

conventional ZSL

bobcat

horse

dolphin
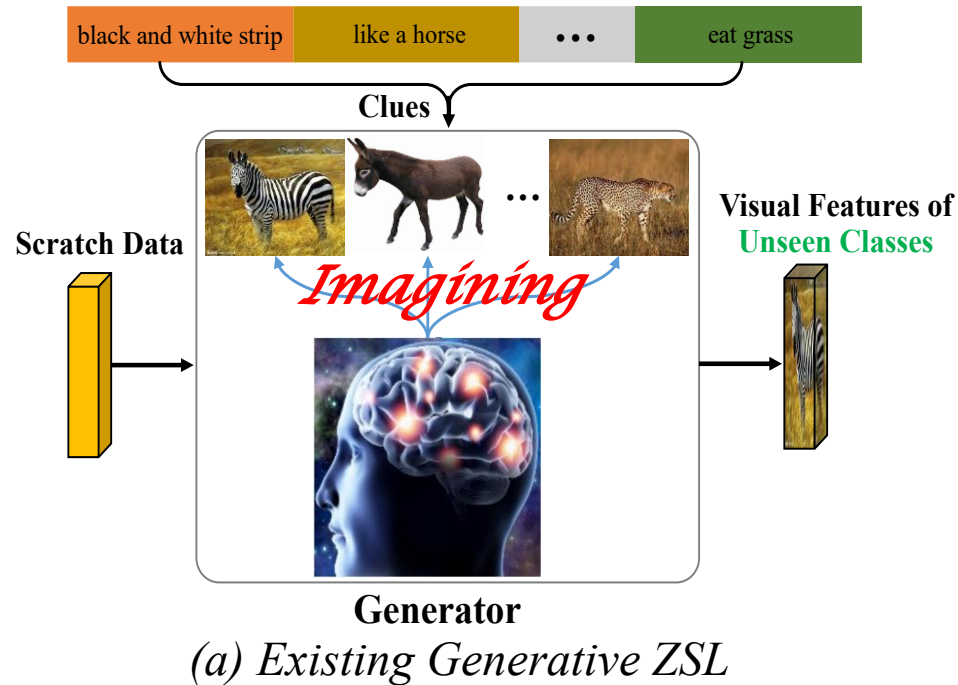
generalized ZSL

$$Y^{tr}$$

$$Y^{te}$$

$$Y^{tr} \bigcap Y^{te}$$
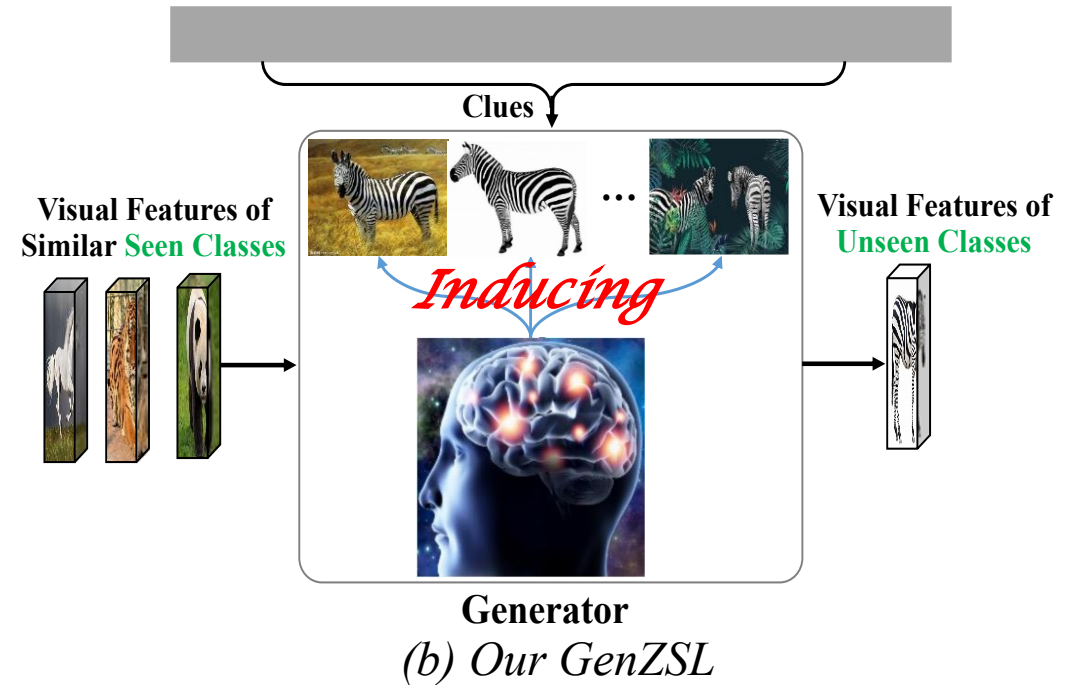
Contents

- Zero-Shot Learning

- Motivation

- The Proposed GenZSL

- Advantages

Motivation

Strong Class Semantic Vector of **Unseen Classes** Annotated by Experts

| black and white strip | like a horse | ... | eat grass |

Clues

Scratch Data

Imagining

Visual Features of **Unseen Classes**

Generator

(a) Existing Generative ZSL

Weak Class Semantic Vector of **Unseen Classes** Extracted by CLIP

Clues

Visual Features of Similar **Seen Classes**

Inducing

Visual Features of **Unseen Classes**
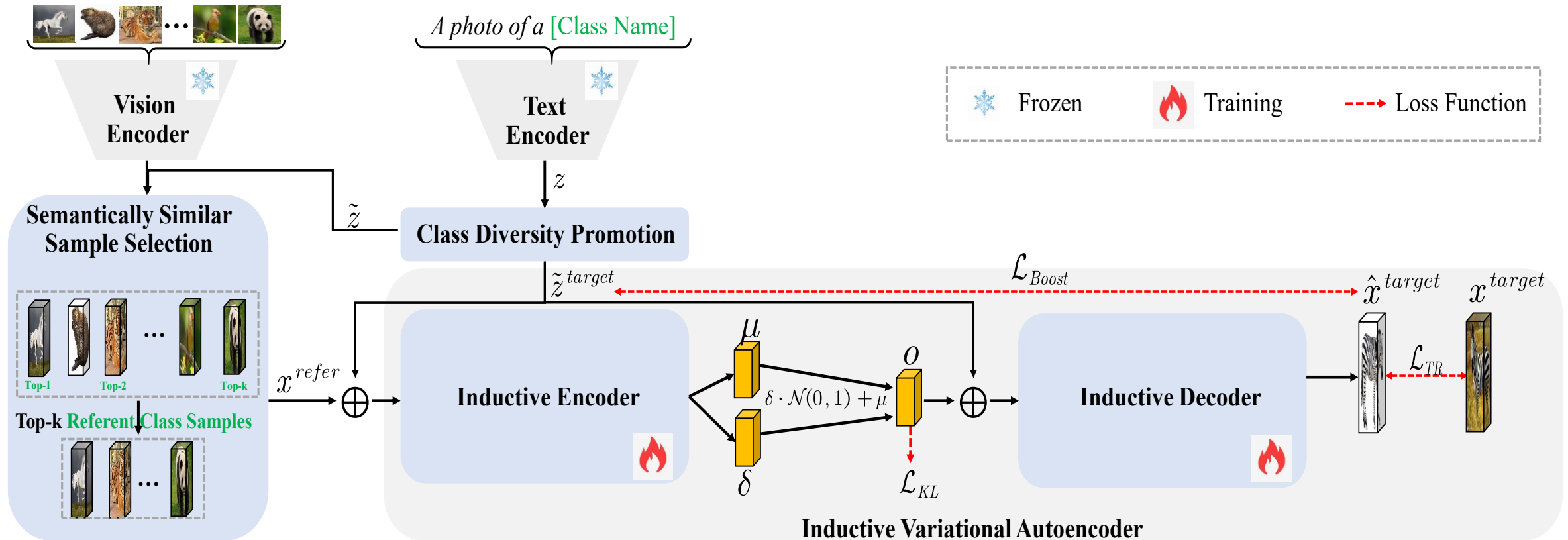
Generator

(b) Our GenZSL

- How to mimic human-level concept learning to inducting new class samples from similar seen classes?

- How to eliminate the dependence on expert-annotated strong class semantic vectors?

Contents

- Zero-Shot Learning

- Motivation

- The Proposed GenZSL

- Advantages

The Proposed GenZSL

Pipeline of our GenZSL

Class Diversity Promotion

- Singular value decomposition of class semantic vectors

$$Z = [z^1, z^2, \cdots, z^C],$$

$$U, S, V = svd(Z).$$

$U = [e^1, e^2, \cdots, e^C]$ is the orthonormal basis.

- Removing the major component e^1

$$U' = [e^2, e^3, \cdots, e^C],$$

$$P = U'U'^T,$$

$$\tilde{Z} = P \cdot Z = \{\tilde{z}^1, \tilde{z}^2, \cdots, \tilde{z}^C\}$$

Semantically Similar Sample Selection

- Selecting similar samples from seen classes using the cosine similarity

$$c^{refer} = arg \max_{topk(c^s)} \frac{\tilde{z}^{target} \times \tilde{z}^{c^s}}{\| \tilde{z}^{target} \| \cdot \| \tilde{z}^{c^s} \|},$$

$\tilde{z}^{c^s}$ is top-k class semantic vectors closed to target class semantic vectors $\tilde{z}^{target}$
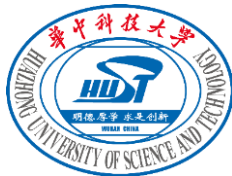
Contents

- Zero-Shot Learning

- Motivation

- The Proposed GenZSL

- Advantages

Advantages

The proposed GenZSL is **simple**, **efficient**, and **effective** to ZSL:

1. It eliminates the dependence on expert-annotated class semantic vectors

2. It proposes a novel inductive variational autoencoder for generative ZSL

3. It demonstrates superiority and potential of our GenZSL with significant efficacy and efficiency over f-VAEGAN, e.g., 24.7% performance gains and more than 60× faster training speed on AWA2

Thank You

Paper Link

Contact Link