



Revisiting Document Representations for Large-Scale Zero-Shot Learning

Jihyung Kil, Wei-Lun Chao
The Ohio State University

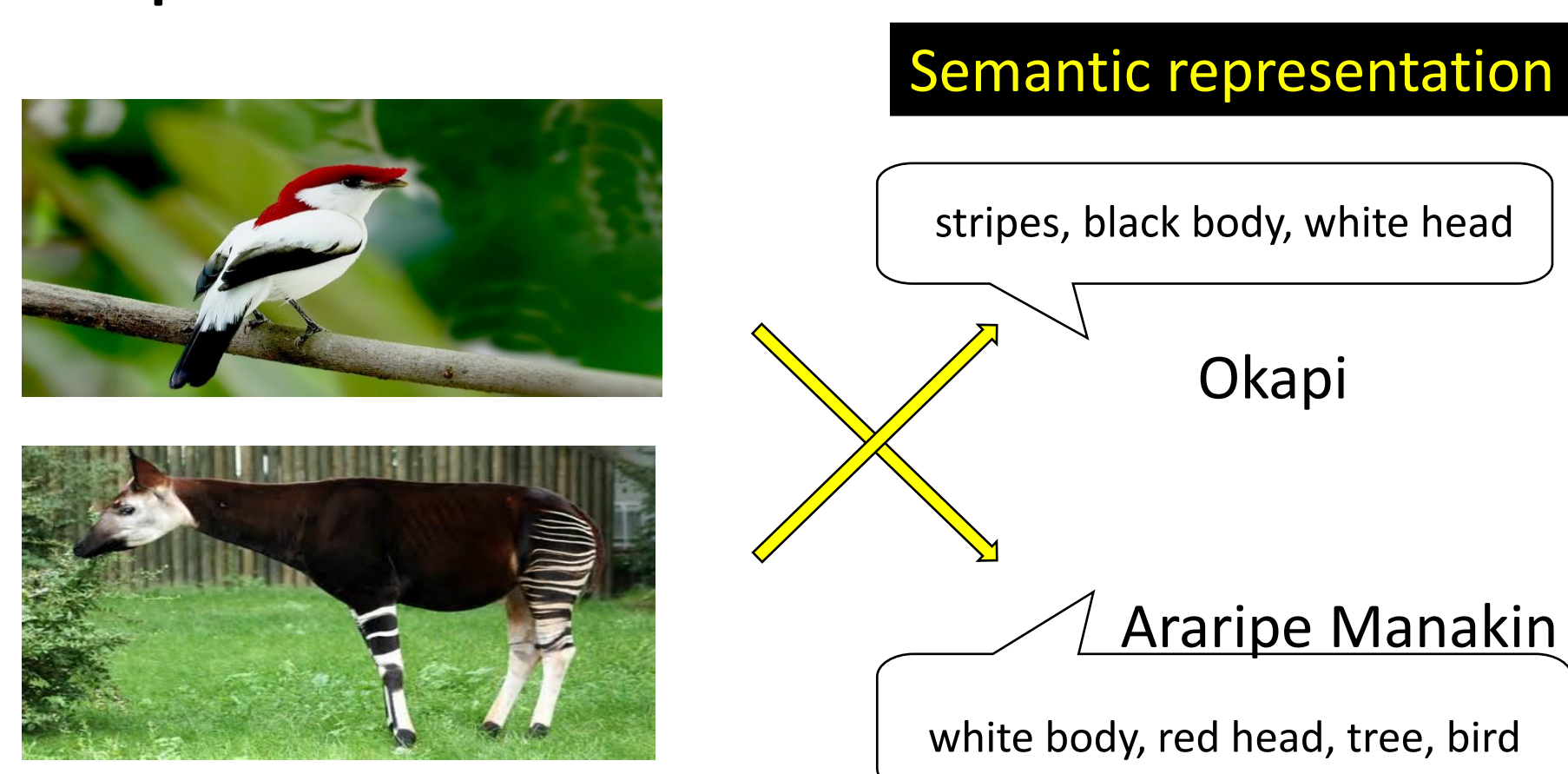
Acknowledgement:
we thank the support of
Ohio Supercomputer Center
and AWS Cloud Credits

Highlights

- Design high-quality semantic representations for Zero-Shot Learning (ZSL) by extracting visual sentences from Wikipedia documents
- Develop a simple semi-automatic mechanism for visual sentence extraction
- Propose a new sentence weighting scheme to distinguish similar classes
- Our approach can be applied to any ZSL algorithms, improving over prior semantic representations by 64% on large-scale ZSL with more than 10,000 unseen classes

Introduction

- Zero-Shot Learning (ZSL) recognizes unseen objects using semantic representations



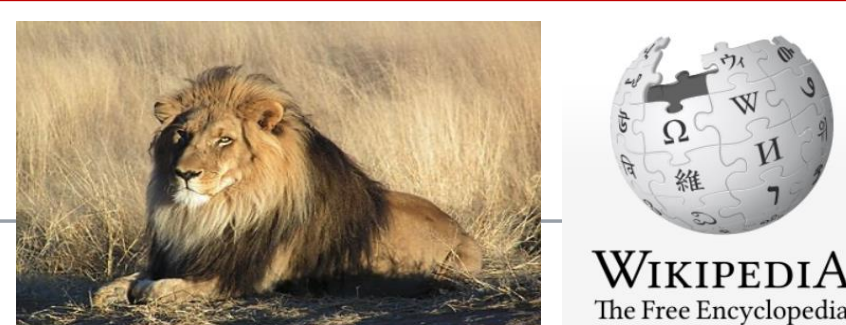
- Popular semantic representations
 - Manually labeled visual attributes [1]
 - Aligned with visual content but not scalable
 - Automatically extracted word vectors of class names [2]
 - Scalable but not aligned with visual contents

For large-scale ZSL problems with many unseen classes, Can we develop semantic representations as powerful as visual attributes without significant human effort?

Visual Section Selection

Lion

From Wikipedia, the free encyclopedia



Contents [hide]

- 1 Etymology
- 2 Taxonomy
 - 2.1 Subspecies
 - 2.2 Fossil records
 - 2.3 Evolution
 - 2.4 Hybrids
- 3 Description
 - 3.1 Size
 - 3.2 Mane
 - 3.3 Colour variation
- 4 Distribution and habitat
 - 4.1 Historical range
- 5 Behaviour and ecology
 - 5.1 Group organisation
 - 5.2 Hunting and diet
 - 5.2.1 Predator competition
 - 5.3 Reproduction and life cycle
 - 5.4 Health
 - 5.5 Communication

Description

The lion is a muscular, deep-chested cat with a short, rounded head, a reduced neck and round ears. Its fur varies in colour from light buff to silvery grey, yellowish red and dark brown. The colours of the underparts are generally lighter. A new-born lion has dark spots, which fade as the cub reaches adulthood, although faint spots often may still be seen on the legs and underparts. The lion is the only member of the cat family that displays obvious sexual dimorphism. Males have broader heads and a prominent mane that grows downwards and backwards covering most of the head, neck, shoulders, and chest. The mane is typically brownish and tinged with yellow, rust and black hairs.^{[47][48]}

Mane

The male lion's mane is the most recognisable feature of the species.^[14] It may have evolved around 320,000–190,000 years ago.^[61] It starts growing when lions are about a year old. Mane colour varies and darkens with age; research shows its colour and size are influenced by environmental factors such as average ambient temperature. Mane length apparently signals fighting success in male–male relationships;

Taxonomy

Felis leo was the scientific name used by Carl Linnaeus in 1758, who described the lion in his work *Systema Naturae*.^[3] The genus name *Panthera* was coined by Lorenz Oken in 1816.^[13] Between the mid-18th and mid-20th centuries, 26 lion specimens were described and proposed as subspecies, of which 11 were recognised as valid in 2005.^[1] They were distinguished mostly by the size and colour of their manes and skins.^[14]

Visual sections

Characteristics, Description, Appearance, Habitat, Diet, Construction and Mechanics, Materials for utensil, Design for appliance, Furnishings for room, Fabrication, Feature for geological formation, Design, Equipment for sport

Visual Cluster Selection

- Visual sentences share common words or phrases of visual attributes, forming a cluster
- K-means sentence clustering with BERT features

kit-fox: red, tiger: blue

Visual Cluster (Top) / Non-Visual Cluster (Bottom)

It has large ears that help the fox lower its body temperature. It usually has a gray coat, with rusty tones, and a black tip to its tail. It has distinct dark patches around the nose.

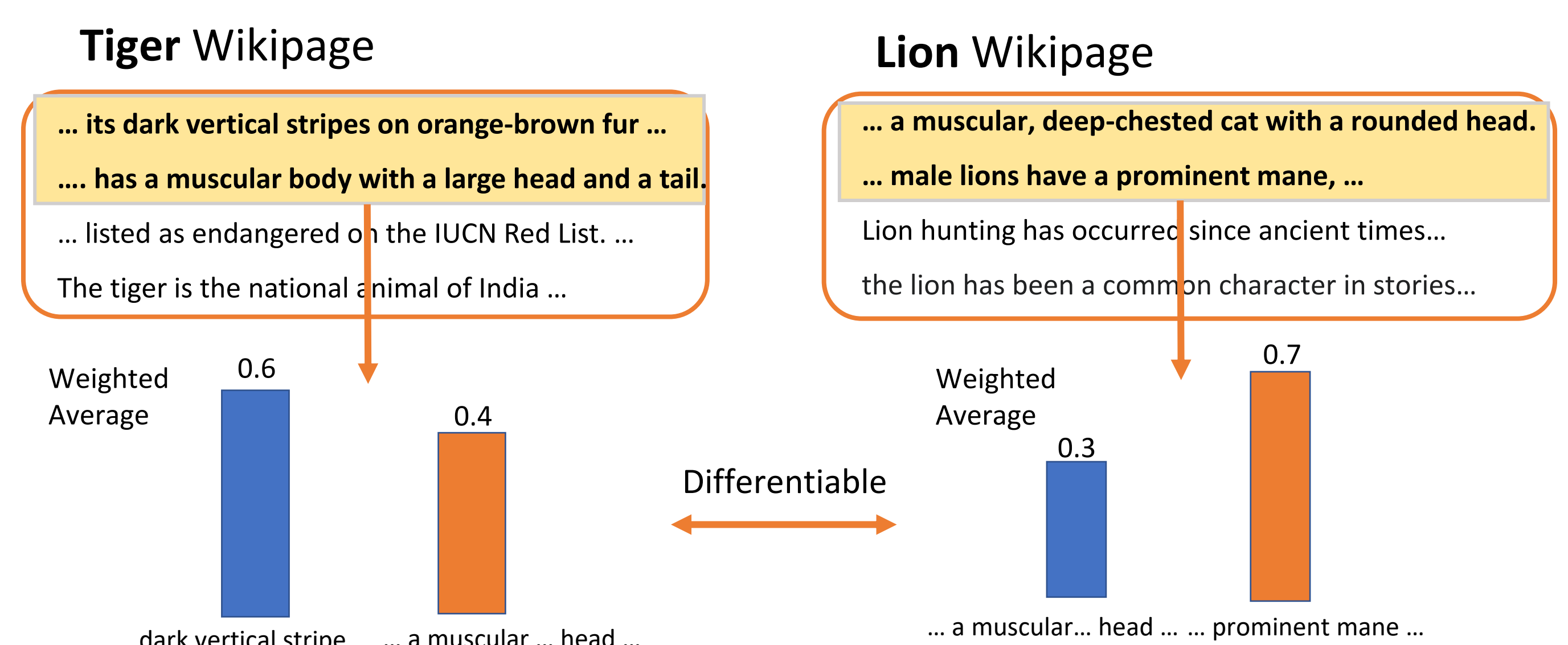
It is most recognisable for its dark vertical stripes on orangish-brown fur. . . . muscular body with powerful forelimbs, a large head and a tail. They have a mane-like heavy growth of fur around the neck and jaws . . .

Male and female kit foxes usually establish monogamous mating . . . The average lifespan of a wild kit fox is 5.5 years.

Tiger mates all year round, but most cubs are born between March . . . The mortality rate of tiger cubs is about 50% in the first two years.

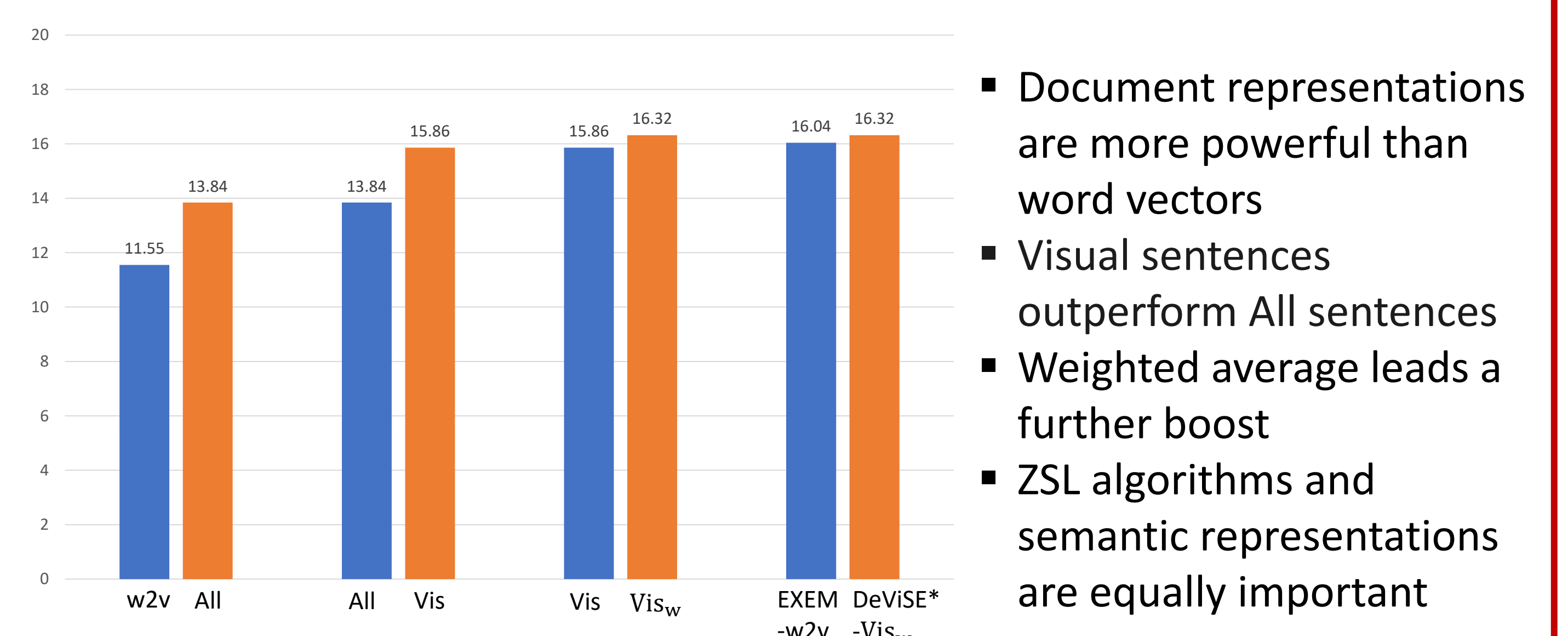
Semantic Representations

- Weighted Average Scheme



Experiments and Analysis

- ImageNet [3]



- Qualitative Results

Image / Label	Type	Top 5 prediction
	Ours	Tiger, Tiger cat, Tabby, Leopard, Jaguar
	w2v	Tiger cat, Tiger, Cougar, Madagascar cat, Standard poodle
	Ours	Sports car, Race car, Garden cart, Minivan, Limousine
	w2v	Jeep, Sports car, ambulance, fire truck, taxi

Conclusion

- Semantic representations are equally important to algorithm design in ZSL
- Existing progress gets stuck due to poor semantic representations

Distinctive Visual sentence extraction on Wikipedia