Recovering the Missing Link: Predicting Class-Attribute Associations for Unsupervised Zero-Shot Learning Supplementary Material

Ziad Al-Halah Makarand Tapaswi Rainer Stiefelhagen
Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany
{ziad.al-halah, makarand.tapaswi, rainer.stiefelhagen}@kit.edu

This document provides additional information about the evaluation of our model.

Obtaining word embeddings We employ a context size of 20, use hierarchical softmax, and learn our skip-gram model using asynchronous stochastic gradient descent and learn a $d = 300$ dimensional word representation. We train our model on the Wikipedia corpus. The corpus includes ~ 3.5B words which yields a 96k word vocabulary that includes high-frequency words and phrases (appearing more than 300 times) and the labels of classes and attributes used in our data sets.

All the class names and most of the attributes used in our study occur naturally in text and their embeddings are obtained directly from the learned dictionary. However, for few attributes like "two_dimensional_boxy" there is no direct match. In such cases, we automatically split the label to the minimum number of constituted phrases that do appear in the dictionary and take the average embedding of these phrases as the representation of the attribute, *i.e.* $\mathrm{v}(two\_dimensional\_boxy) = \frac{1}{2}(\,\mathrm{v}(two\_dimensional) + \mathrm{v}(boxy)\,)$

Semantic relations For the semantic relation model (SR), we defined a set of 9 relations in AwA. Table 1 shows the complete list of the relations used and their associated attributes.

Associations prediction Fig. 1 shows the association accuracy for each individual attribute in AwA and aPaY.

Model hyperparameters In Fig. 2, we present the performance of CAAP during cross validation with respect to hyperparameters L and λ. The figure shows the mean average precision of association prediction in AwA data set. We notice that λ has less impact on the model compared to L.

Qualitative results Fig. 3 shows the top 5 most confident zero-shot classifications of our model for each of the unseen classes in AwA. Correct predictions are marked with green while the wrong ones in red. Most of the confident classifications are correct. Table 2 shows the top 5 most confident associations for all the unseen classes in AwA.

Relation	Attribute
has_color	black, white, blue, brown, gray, orange, red, yellow
has_pattern	patches, spots, stripes, furry, hairless, toughskin
has_shape	big, small, bulbous, lean, muscle, bipedal, quadrapedal
has_part	flippers, hands, hooves, pads, paws, longleg, longneck, tail, chewteeth, meatteeth, buckteeth, strainteeth, horns, claws, tusks
movement_type	flys, hops, swims, tunnels, walks, fast, slow, agility
food_type	fish, meat, plankton, vegetation, insects
feeding_style	forager, grazer, hunter, scavenger, skimmer, stalker
lives_in	newworld, oldworld, arctic, coastal, nestspot, desert, bush, plains, forest, fields, jungle, mountains, ocean, ground, water, tree, cave
behavior	fierce, timid, smart, group, solitary, domestic, strong, weak, active, inactive, nocturnal, hibernate, smelly

Table 1: The set of defined semantic relations in AwA and their respective attributes.

Unseen Category	Top Associations Positive	Negative
persian_cat	tail, fast, paws, *active*, furry	orange, yellow, horns, tusks, desert
hippopotamus	strong, *group*, big, walks, ground	claws, flys, red, nocturnal, weak
leopard	fast, lean, oldworld, active, tail	tusks, water, arctic, plankton, weak
humpback_whale	fast, ocean, water, group, fish	red, weak, tunnels, nocturnal, plains
seal	fast, *meatteeth*, *bulbous*, big, toughskin	grazer, tunnels, longleg, hooves, longneck
chimpanzee	walks, group, fast, chewteeth, active	arctic, flippers, red, plankton, strainteeth
rat	furry, active, *chewteeth*, newworld, fast	plankton, yellow, orange, horns, desert
giant_panda	*fast*, walks, *active*, quadrapedal, strong	*domestic*, weak, strainteeth, desert, flys
pig	ground, timid, white, chewteeth, quadrapedal	cave, plankton, orange, desert, yellow
raccoon	fast, newworld, quadrapedal, furry, active	bush, hands, longneck, tusks, desert

Table 2: Predicted class-attribute associations for all unseen classes in AwA. Wrong associations are highlighted in gray and italic.

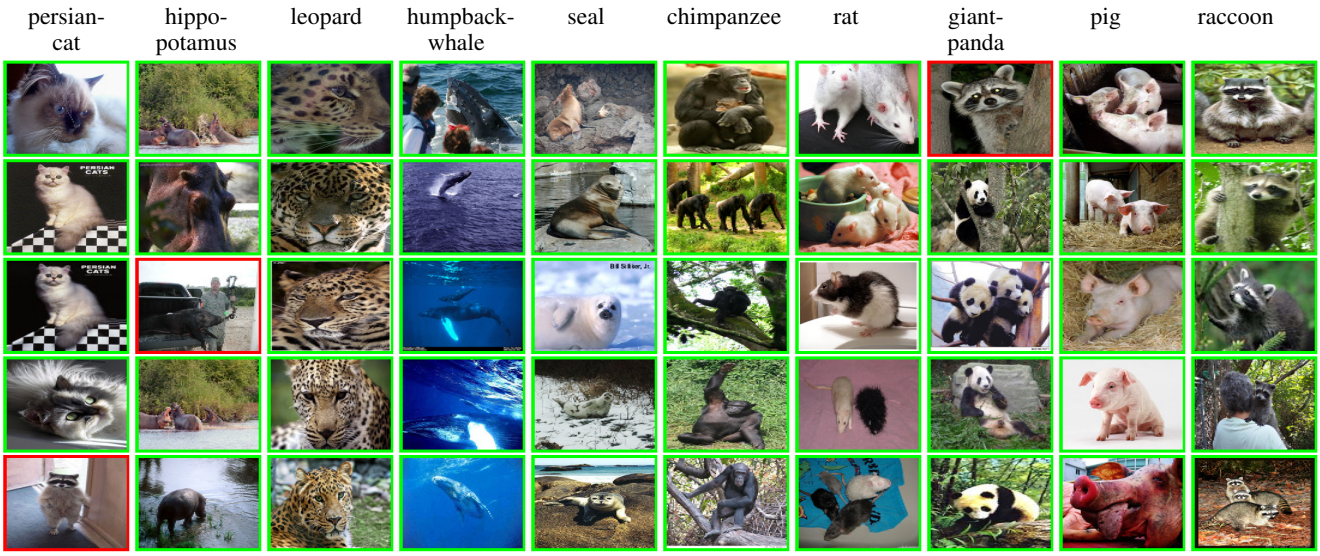
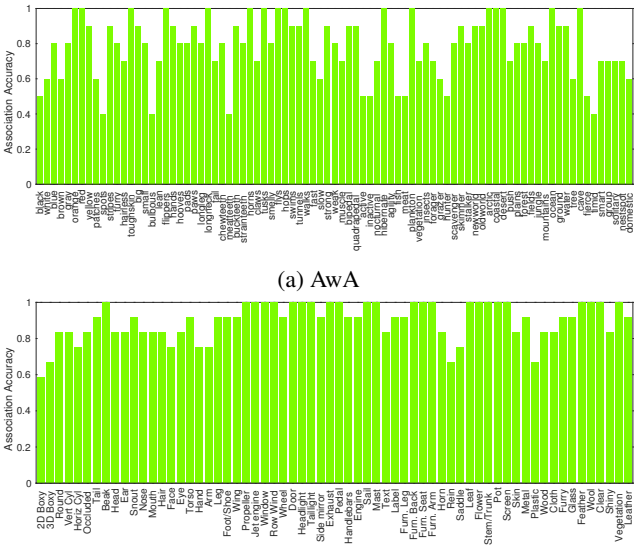


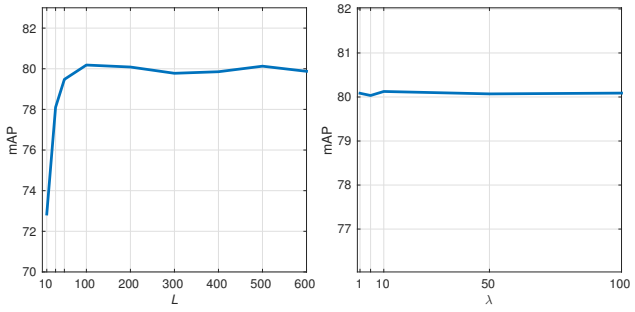
Figure 3: The top 5 ranked results of our CAAP model for each unseen class in AwA. (Best viewed in color)



(a) AwA

(b) aPaY

Figure 1: Accuracy of predicting individual attribute associations over the unseen classes.



(a) (b)

Figure 2: The cross validation performance of our model in AwA when (a) varying number of latent factors L while $\lambda = 1$ and (b) varying the value of λ while $L = 200$.