Progress Overview

Trying to run Karpathy code on Paris (working)

*Deep Visual-Semantic Alignments for Generating Image Descriptions,* Karpathy & Fei-Fei

Extracted the flickr30k entities

*Flickr30k Entities: Collecting Region-to-Phrase Correspondences for Richer Image-to-Sentence Models*, Plummer et al.

LDA on sentences of flickr30k

*Aligning where to see and what to tell: image caption with region-based attention and scene factorization*, Jin et al.

Based on the LDA distributions for the Flickr30k training set, we trained a neural network to predict the distribution based on a picture representation. We found the results very interesting.

Basic FSMN implementation

*Feedforward Sequential Memory Neural Networks without Recurrent Feedback*, Zhang et al.