

Dataset: Flickr 80k datasets with photo descriptions

(<https://forms.illinois.edu/sec/1713398>)

Download glove.6B.200d.txt from UCB for word semantics embedding

Task: train a model for generating a photo description (caption) from input photos

Approach: use Densenet (a pretrained convolutional neural network) to extract the photo features, and then use LSTM for the simple NLP task for generating a short caption for the photos

Example output:

