Dataset: Flickr 80k datasets with photo descriptions (<https://forms.illinois.edu/sec/1713398>)

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 the short caption for the photos

Example output:

![A screenshot of a cell phone

Description automatically generated]()

![A screenshot of a cell phone

Description automatically generated]()