Adam Zucker
EE/CS '17
CPSC 490
03/09/17

Web Application for Generating Image Descriptions Using Natural Language Processing and Computer Vision Techniques

Description:

For my final project, I am building a web application that builds upon existing computer vision and natural language processing techniques to generate descriptions for images a user uploads.

What I have done:

So far, I have fully installed and set up all of the front-end and back-end dependencies needed for my web app on the dedicated Zoo machine, Tangra. The web application is now fully functional, but currently has only a single skeleton view. That being said, I foresee this being a single-page web app, so there should not be too much more work required for the web application side of things.

Additionally, I have installed and set up the computer vision and natural language processing software I will need to use to for the caption generation algorithm. These include the NLTK library, MSCoco API and corresponding datasets, and PyTorch (deep learning Python library). I have experimented a bit with the MSCoco dataset, and have been able to successfully load image captions.

I have uploaded all of my work to a private Github repository called image-caption-generation (**https://github.com/adamzu/image-caption-generation**).

What remains to be done:

The first thing I need to do is actually completely implement my single-page web app. I had listed "website interactions mocks/sketches" as one of my deliverables, so I will draw those up as well, before I make the app. As said, this part of the project should not require too much more time.

The main thing I still need to do is implement my caption generation algorithm. I have read some papers on MSCoco and watched some tutorials on PyTorch, so I feel like I know how to proceed; now it is time to get to testing out some different caption generation approaches.

Finally, I need to write up my report that describes what my app is and how it works. I plan on working on this report concurrently with my caption generation algorithm, and cannot imagine it will be particularly challenging to complete.

Issues:

So far, I have not encountered too many significant snags, but this is probably mainly due to the fact that I have not done much with actual caption generation algorithm. My issues thus far have mainly been related to installing dependencies for my web application, but they have all been resolved.