

Project Title: Image Captioning

Project Objective

With the rapid development of artificial intelligence in recent years, image captioning has gradually drawn the attention of many researchers in the field of artificial intelligence and has become an interesting and challenging task. Image captioning, which automatically generates natural language descriptions based on the content observed in an image, is an important part of scene understanding, which combines computer vision and natural language processing knowledge. The use of image captions is widespread and important, for example, in the realization of human-computer interaction.

Read: <https://www.hindawi.com/journals/cin/2020/3062706/>

The problem introduces a captioning task, which necessitates a computer vision system that can both localize and describe salient regions in images in natural language. When descriptions consist of a single word, the image captioning task generalizes object detection. Find the correct semantic label for the entire image given a set of images and prior knowledge about the content.

Customer Requirements

1. Build a natural language image captioning model *not* based on *transfer learning*
2. The model accepts an image as input and returns Natural language description of the input image
3. Using the django framework, design an integrated interface to display the results of the model results from a user input
4. As a user, I want the model returned descriptive output to be used as a search input on the web, and the corresponding images results displayed
5. Data Sources: [primary](#) (feel free to sample the dataset with clear justification); secondary- MS COCO dataset, Flickr30k dataset

Deliverables

- Web application source code (GitHub - implement Gitflow with reasonable number of commits including a well-written README file)
- A working django UI deployment via Heroku
- Final project presentation