Image Captioning

8th September 2018

OVERVIEW

Dealing with Out of Vocabulary (OoV) words and detection of previously unseen objects in an image are the current main challenges in Image captioning. In this project we aim at addressing the challenge to generate captions for images which contain unseen objects. Good amount of work has been done in this field like KGA-CGM, LSTM-C, DCC etc. We would either try to make a hybrid of such methods or build a new model from scratch to get the best possible solution.

GOALS

1. To make an Image Caption Generator which can generate caption(s).
2. To extend the model to generate captions for unseen images also.

RELEVANT PRIOR WORK

1. Describing Natural Images Containing Novel Objects with Knowledge Guided Aissitance - https://arxiv.org/pdf/1710.06303.pdf
2. Incorporating Copying Mechanism in Image Captioning for Learning Novel Objects - https://arxiv.org/pdf/1708.05271.pdf.
3. Deep Compositional Captioning: Describing Novel Object Categories without Paired Training Data - https://arxiv.org/pdf/1511.05284.pdf

SPECIFICATIONS

The task of image captioning will be divided into two modules one is image based model to extract features and nuances of images and other is language based model which translates the features and objects given by our image based model to a natural sentence. Our image based model would include CNN whereas language based model would use RNN or LSTM.

DATASETS

1. MSCOCO Dataset - http://cocodataset.org/#download

BRIEF TIMELINE

Learning Phase

Reading relevant works on Knowledge Guided Assistance (KGA) and Copying mechanism which are potential solutions to address the challenge.

Implementation Phase

Implementing the above two models using libraries like TensorFlow, PyTorch, etc.

Improvisation Phase

Improving the accuracy of the model by further trying to come up with better modifications of CNN and RNN architecture.

OUR TEAM

Pragya Jain - 160492 - pjjain@iitk.ac.in
Rahul B S - 160540 - bsrahul@iitk.ac.in
Rahul Saxena - 160549 - rahulx@iitk.ac.in
Subham Kumar - 160707 - subhamkr@iitk.ac.in
Tushar Gurjar - 160750 - tushargr@iitk.ac.in
Vinayak Trivedi - 160790 - vinayakt@iitk.ac.in