Dynamic Lexicon Generation for Natural Scene Images

Joyneel Misra
Sarthak Sharma
Shubham Rathore

Mentor: Yash Patel

Project Outline

- Existing studies aim at tackling text reading from natural scene images using customised lexicons
- The size and quality of the lexicons have a huge effect on these systems
- The project aims at dynamically generating contextualized lexicons for scene images exploiting the rich visual information in a dataset consisting of images and textual information associated with them

Our understanding of the project

- A method to generate context based lexicons using the rich visual information
- Major assumption: some words in the natural scene image are correlated directly to the objects appearing in the image
- There is an evidence of correlation between visual and textual information
- Topic models are to be used to exploit the relation between visual and textual information for image annotation.

Topic Model

LDA: A topic model which helps in finding latent topics out of a set of documents and predicts probability distribution of words belonging to a topic, for each topic.

CNN:Incorporating this topic model with a CNN we are able to learn the probability distribution as we did in LDA but the only difference being is the usage of visual data

Input for training **CNN** -------> Image + Topic probability per image[from **LDA**]

Method/Scope

Step 1: We learn an LDA topic model on a text corpus associated with the image dataset

Step 2: We train a deep CNN model to generate LDA's topic probability directly from image pixels

Step 3: Generated topic probabilities are used from the LDA models or from the CNN , along with the word-probabilities from the learned LDA model to re-rank the words of a given dictionary

Frameworks

For **LDA** : Gensim

For **CNN** :Tensor Flow (Fine tuning Inception_v3 model)

Project extension ideas :

- Generation of an image, given a textual description
- Exploiting similarity in images for the generation of more scalable and efficient lexicon
- Designing an algorithm that merges the two dictionaries in an efficient way