Deep Learning for Image Captioning

By
Siddesh Pillai (srp4698@rit.edu)

Advisor : Professor Jeremy Brown (jsb@cs.rit.edu)
Colloquium Advisor : Professor Leon Reznik (lr@cs.rit.edu)

Agenda

- Reminder
- Updates of Milestone-2 tasks
- Challenges Faced
- Next Steps

Reminder

- **What is Deep Learning?**
- A branch of machine learning based on set of algorithms that attempts to model high level abstractions in data by using multiple processing layers with complex structures [1]
- **The Problem Statement** – Image Captioning
- **Base Implementation** – Deep Visual Semantic Alignments for Generating Image Descriptions [2]

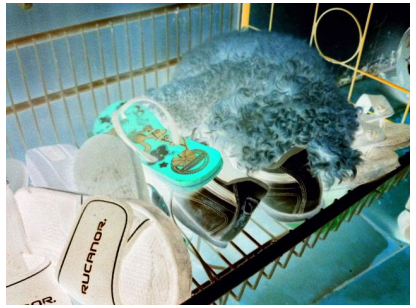
Milestone – 2 tasks (1/2)

- Apply Filters [3] :

(original image)

A dog is sitting on a skateboard in a room

Complement



A cat is sitting on a chair with a stuffed animal

High light Intensity

A cat is sitting on the floor next to a pair of scissors

Low light Intensity

A cat sitting on top of a pair of shoes

Gaussian

A close up of a person holding a cell phone

Milestone – 2 tasks (2/2)

- Scale Image Resolutions [4] :
  - High Resolution (1600 x 1600 pix)
  - Medium Resolution (720 x 720 pix)
  - Low Resolution (120 x 120 pix)
- Extract image meta-data [5]
- Model the classifier based on filters and resolution
- Started working Android app – Image labelling (which is part of milestone 3)

Challenges

- Applying image filters was carried using Open CV library in python tackled lots of dependency issues and then moved to Matlab
- Extracted the image meta-data to find out the homogeneity of the images

Next Steps

- Analyze measured results and understand how deep captioning networks behaves to computationally altered images
  - Impact of color variance
  - Impact of spatial arrangements
- Complete working on android app
- Start documenting the project work
- Start working on poster design

References

1. Deep Learning https://en.wikipedia.org/wiki/Deep_learning
2. Andrej Karpathy and L. Fei-Fei. Deep Visual-Semantic Alignments for Generating Image Descriptions, 2012
3. Image Filters https://github.com/sidd4698/capstone-project/tree/master/image-filter
4. Image Resolutions https://github.com/sidd4698/capstone-project/tree/master/image-res
5. Extract Image meta-data https://github.com/sidd4698/capstone-project/tree/master/image-metadata

Thank you