Z.H.C.E.T., Aligarh Muslim University

B. Tech Project Part - I

(COC 4980)

Visual QnA System

Week Ending 6th Sep

Under the Supervision of **Prof. Mohammad Sarosh Umar**

Ravi Sahni

17 COB 085
GJ 7718

Tanuj Singhal

17 COB 033
GI 1585

Objective:

We intend to implement the task of free-form and open-ended Visual Question Answering.

Abstract:

Given an image and a natural language question about the image, the task is to provide an accurate natural language answer. Starting from basic images-less features like small shapes and moving on to real life images with huge feature set.

Progress:

We spent the last week searching for relevant research papers to refer to. These papers were:

[1] Agrawal, A., Lu, J., Antol, S., Mitchell, M., Zitnick, C., Parikh, D. and Batra, D., 2016. VQA: Visual Question Answering. *International Journal of Computer Vision*, [online] 123(1), pp.4-31. Available: <https://arxiv.org/pdf/1505.00468.pdf>.

[2] Goyal, Y., Khot, T., Agrawal, A., Summers-Stay, D., Batra, D. and Parikh, D., 2018. Making the V in VQA Matter: Elevating the Role of Image Understanding in Visual Question Answering. *International Journal of Computer Vision*, [online] 127(4), pp.398-414. Available: <https://arxiv.org/pdf/1612.00837.pdf>.

This week we tried to synchronize and assimilate from whatever insights both of us could gain from the papers listed above. We found the topic to be quite new for us with mentions of a lot of new terminologies, so we believe we will need more time to go through these papers.

Also, we have found some more papers to explore and expand our search for the one we can implement in our resources. These are listed below:

[1] L. Ma, Z. Lu and H. Li, "Learning to Answer Questions From Image Using Convolutional Neural Network", 2015 [Online]. Available: <https://arxiv.org/pdf/1506.00333.pdf>.

[2] M. Ren, R. Kiros, and R. Zemel, "Exploring Models and Data for Image Question Answering", 2015 [Online]. Available: <https://arxiv.org/pdf/1505.02074.pdf>.

****