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

B. Tech Project Part - I
(COC 4980)

Visual QnA System

Week Ending 30th Aug

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.

References:

We spent most of the week searching for relevant research papers to refer to. Some of the research papers we found are listed below:

[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 at: <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 at: <https://arxiv.org/pdf/1612.00837.pdf>.

We plan to go through these papers thoroughly and see if the techniques discussed in these papers are implementable within our limited resources if not then we plan to extend our search for other papers.

****