

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

## **B. Tech Project Part - I**

**(COC 4980)**

# **Visual QnA System**

***Week Ending 13th 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:**

In continuation from last week, we started to explore and download the datasets available for our task and got stuck as the two famous datasets for the VQA task were far too big to handle on our local systems so we decided that we will try to find some form of substitute for the same later, and in the meanwhile go on to understand the terminologies that are being talked about in the papers. We will get started with basics of natural language processing this week and then continue on to algorithms using which it can be applied in accordance with visual deep learning architectures to fit the required tasks.

\*\*\*\*\*