

## **Image Captioning and Android App for Visually Impaired persons**

**Mentor: Dr. Mahima Kaushik**

**Students - Rddhima Raghunand, Shobhit Maheshwari**

Computer vision has become ubiquitous in our society, with applications in several fields. In this project, we focussed on one of the visual recognition facets of computer vision, i.e image captioning. The problem of generating language descriptions for visual data has been studied from a long time but in the field of videos. In the recent few years emphasis has lead on still image description with natural text. Due to the recent advancements in the field of object detection, the task of scene description in an image has become easier.

The aim of the project is to improve the current standard architectures available for the task of Image Captioning and integrate the model with an Android Application, so as to make an app that can work as real-time scene descriptor. The app can be of great utility to Visually Impaired persons, since it can help them navigate, browse and perform other day-to-day activities. In this project, we are trying to improvise on the deepnet, which was based on a combination of VGG16 encoder and an LSTM decoder.

