

PROJECT PROPOSAL

ABSTRACT

We would like to take up **Semantic cognition model from images with question answering** as listed on the Final Projects Ideas Page.

We propose to build a Visual Question Answering network inspired by [1]. This will use an LSTM to process the natural language query. Images will be processed by a convolutional neural network. The image and question embeddings will then be combined as proposed in [1] to answer the query. We aim to train the network on at least the R&M model facts and study the **dynamics of degradation** when noise is added to the network.

Image data will be accumulated from ImageNet or Kaggle and the Question & Answer set will be modelled manually by us since for a start we will only try to get the binary output (Yes or No).

REFERENCE

[1] - Agrawal, Aishwarya, et al. "Vqa: Visual question answering." *International Journal of Computer Vision* 123.1 (2017): 4-31.

TEAM MEMBERS:

1. Kavya Sree Bhagavatula - ksb456
2. Poornima Haridas - ph1391