

## Hierarchical Model for Cross Modal Retrieval

In this project, we will try to explore cross modal retrieval (CMR) where the task is to retrieve an image given its captions or retrieve a caption given an image. The main idea of the project is to create CMR models that can perform retrieval across different datasets by creating hierarchical embeddings. We will try to answer the hypothesis that learning concepts hierarchically will result in better domain adaptation and generalization.

More specifically, we will train models on Flickr30k while testing on MSCOCO (or vice versa) by learning hierarchical concepts present in the captions. In this project, you will get familiar with NLP tools and data such as WordNet, ConceptNet, FrameNet, spacy, gensim, etc.

