

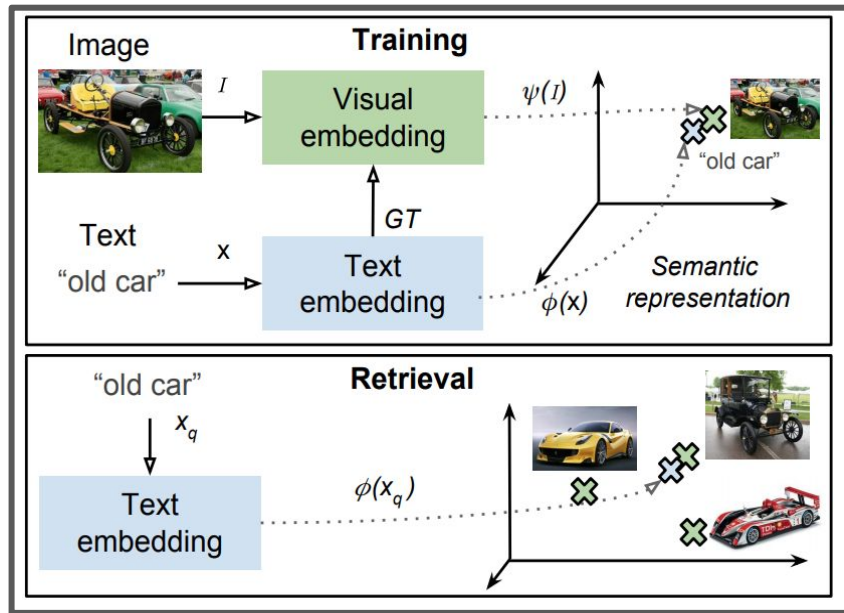
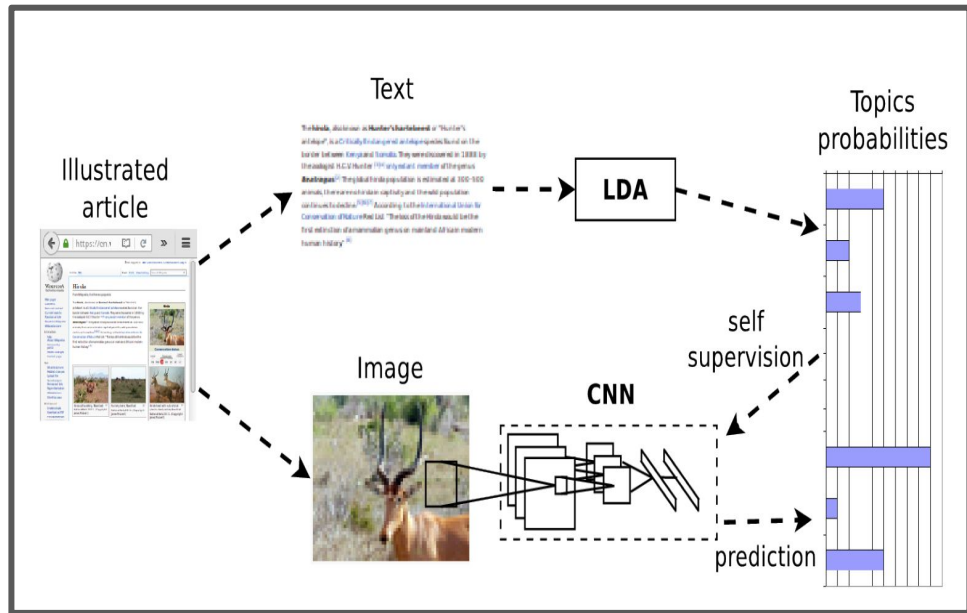
Self-Supervised Learning on Wikipedia Articles for Multimodal Retrieval

Nisarg Thakkar (nmt324)

Srishti Grover (sg5783)

Sree Lakshmi Addepalli (sla410)

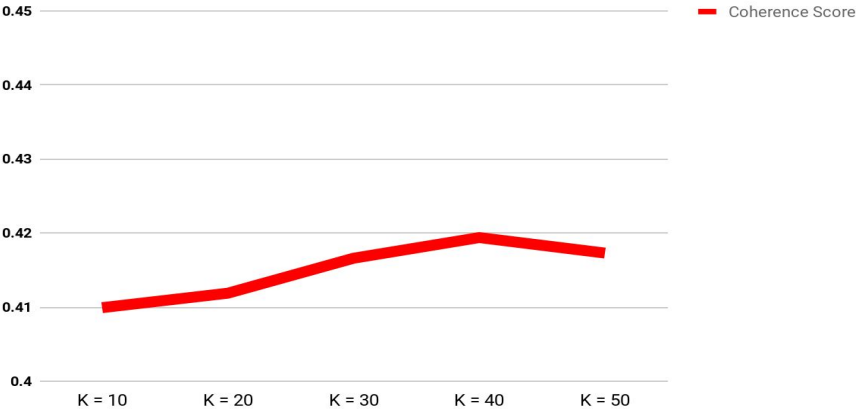
Architecture



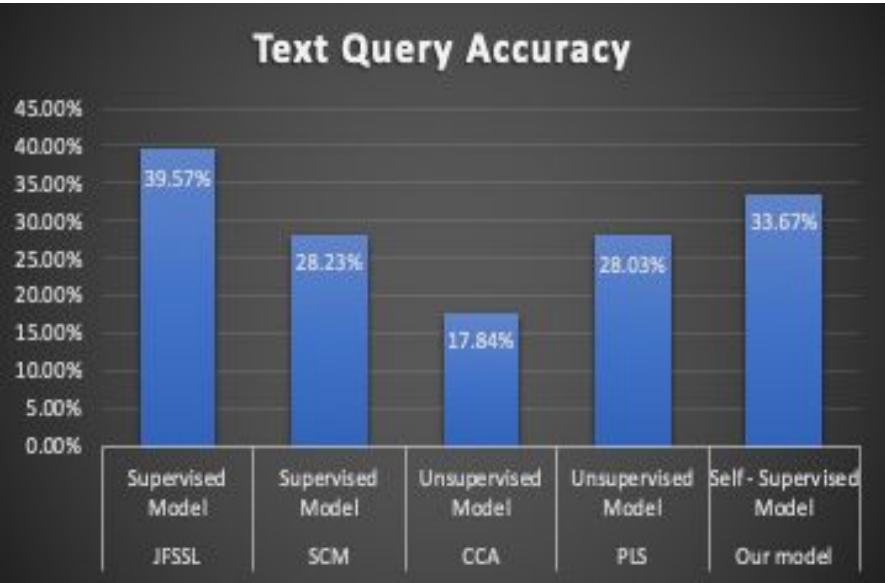
Problem:

1. End-to-end training of deep architectures for new computer vision problems.
2. Would require Imagenet-scale datasets.
3. It is time consuming to generate labels manually.

Number of Classes For LDA according to Coherence Score



Model	Type	Text Query Accuracy
CCA	Unsupervised Model	17.84%
PLS	Unsupervised Model	28.03%
Our model	Self - Supervised Model	33.67%



Model	Type	Text Query Accuracy
JFSSL	Supervised Model	39.57%
SCM	Supervised Model	28.23%