Instance Based Image Retrieval

Project Review

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

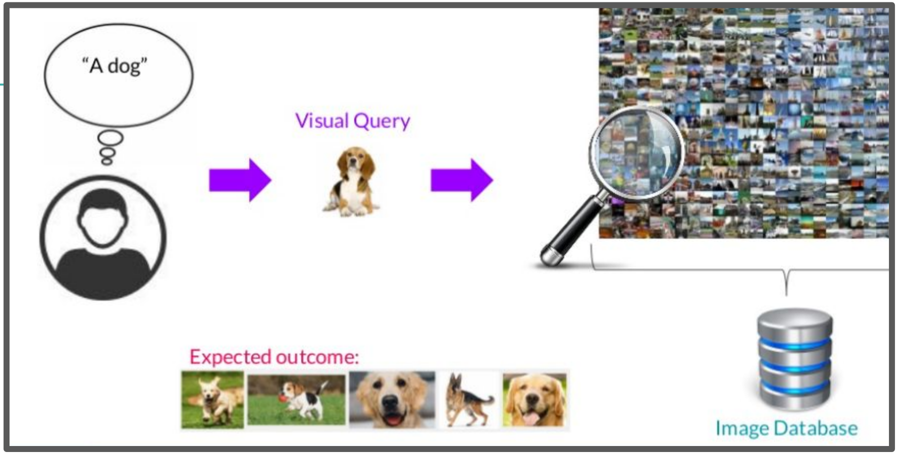
Problem Statement

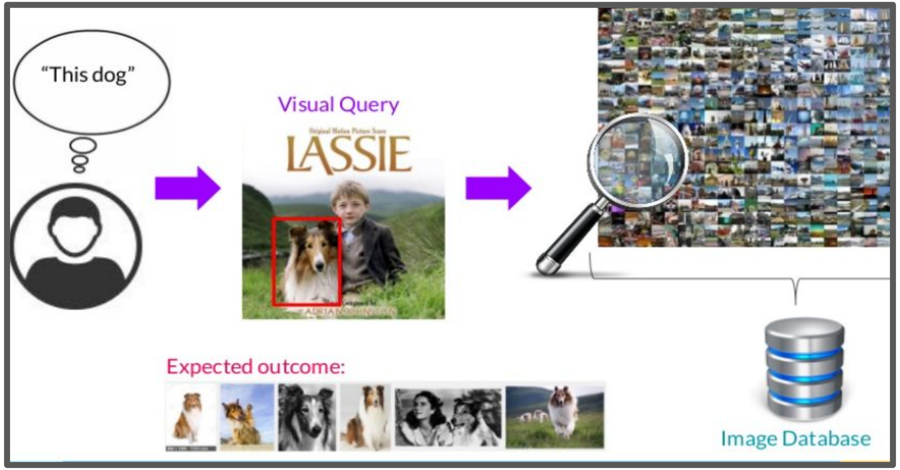
Problem Statement

1. Solving the problem of Instance based Image retrieval.

1. The task is to retrieve images from the database based upon the object in the target image.

1. Useful in the field of medical sciences, astronomy, security, autonomous vehicles among others





Literature Review

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

Literature Review

- In the past decade, image retrieval types:cross-modal retrieval, sketch based retrieval, multi-label retrieval, instance retrieval, object retrieval, semantic retrieval, fine-grained retrieval.
- Earlier techniques uses hand-crafted features for image matching but now shifting towards deep learning-based features.
- some Researchers have used techniques like SIFT and CNN based features for instance based image retrieval.

- SIFT based features do not perform well on specific object retrieval,
  - CNN-based techniques require a sufficient amount of data for training.

- Another baseline model in instance retrieval is R-MAC, but it considers fixed spatial pooling, leading to higher computations.
  - To overcome this extend the R-MAC approach is used. Extend R-MAC produced a global image representation by aggregating the activation features.

Limitation in current approaches

- A breadth and depth analysis on extracting better features from bounding box.
- Most of the models use only cosine similarity.
- Analysis over variable datasets.
- SIFT features used which fail to capture relevant features.

Methodology

Methodology

Training Steps

1.Extraction of Features from Images using Feature extraction techniques.

2.Apply K-Means Clustering on on all images features and select P clusters as Visual Words.

3.Assign closest visual word to each pixel of dataset images.

4.Represent each image as Bag of Word of Visual Words.

Methodology

Retrieval

1.Apply Same pre processing to convert query image into BoW vector.

2.Apply Cosine Similarity with BoW of dataset images.

3.Select Top K images and apply local search by windowing retrieved images into patches of some aspect ratio. (Different methods to get Top results)

4.Returned reranked results

ML Based Baseline

ML Based Baseline

- For instance based retrievals we have used the SIFT technique to find the salient features in the image.
- we generate a visual codebook by applying K Mean Clustering on all features.
- This codebook represents each image as Bag of these Visual Words(BoW).

Procedure

- we first retrieve images which are having highest Cosine similarity of BoW with that of query image are highest.

- We now expect to have visually similar images with some noisy data.

- we select top K images and re-rank them by finding cosine similarity within BoW of query image with sliding windows of that of top K images.

- This arranges the image containing the query instance in top rankings

Results

Results

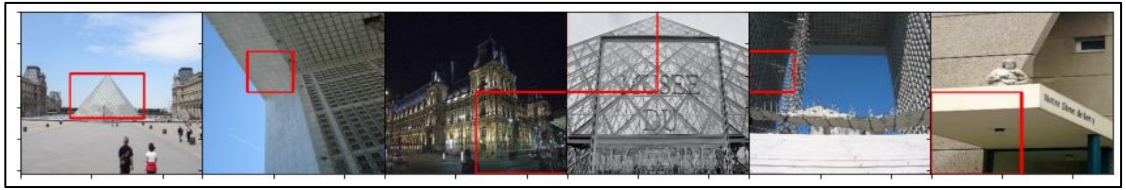
SIMILARITY : COSINE

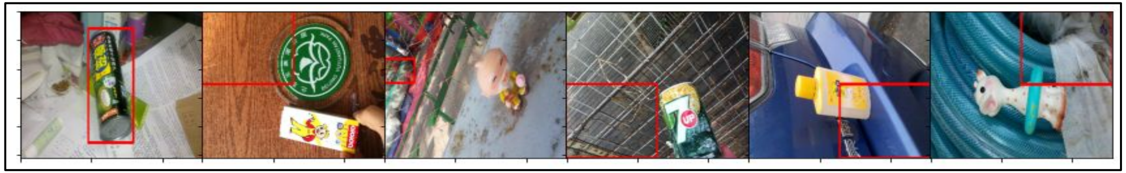
Dataset	MAP	NDCG
Oxford	0.0334	0.1824
Paris	0.0556	0.3883
INSTRE	0.0627	0.2904
Sculpture	0.0654	0.3021

SIMILARITY : WEIGHTED COSINE

Dataset	MAP	NDCG
Oxford	0.0184	0.1667
Paris	0.0719	0.4296
INSTRE	0.0650	0.3123
Sculpture	0.1620	0.5384

Results



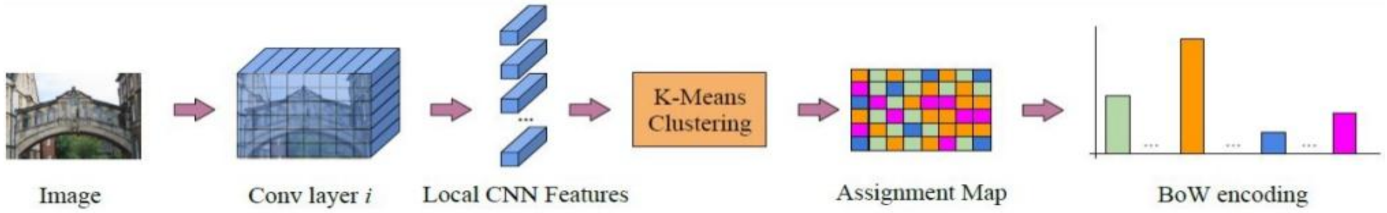


DL Based Model-1

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

DL Based Baseline

- Implementementeded [Bags of Local Convolutional Features for Scalable Instance Search]
- Features Extracted using VGG-16
- Image Representation using BOW model

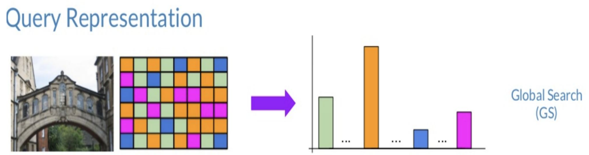


- Similarity matching using cosine similarity
- Global and Local Ranking approach

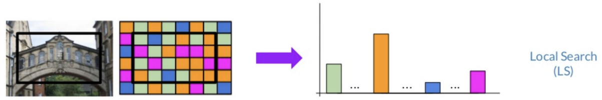
DL Based Baseline

- Global Ranking - Comparison across all images (cosine similarity between Bow of query image and Database image)

  Top M results



- Local Ranking approach - Reranking of Top M results based on query bounding box and window sliding approach



Results

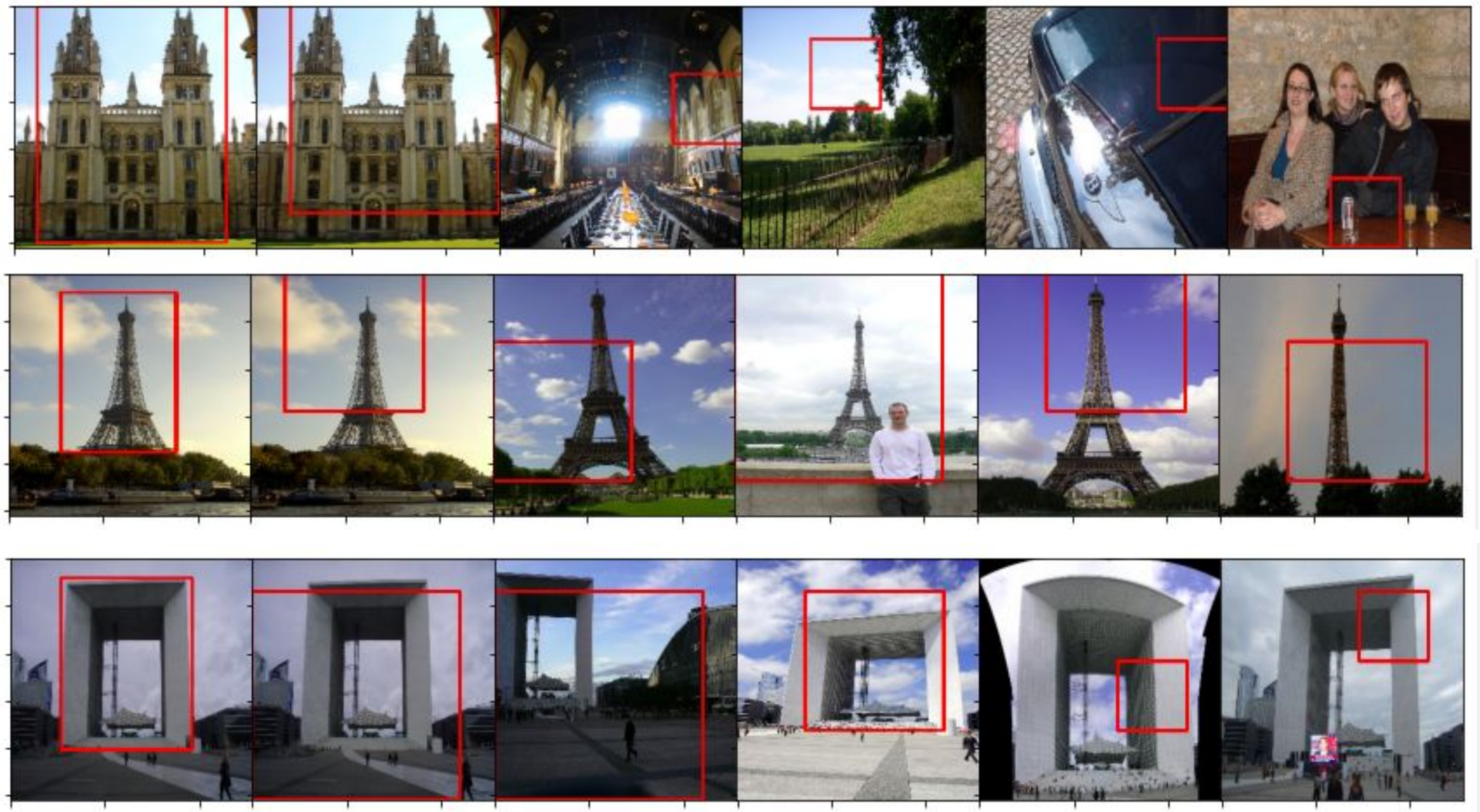
SIMILARITY : COSINE

Dataset	MAP	NDCG
Oxford	0.6963	0.7539
Paris	0.6296	0.8083
INSTRE	0.1862	0.3964
Sculpture	0.0654	0.3021

SIMILARITY : WEIGHTED COSINE

Dataset	MAP	NDCG
Oxford	0.6961	0.7535
Paris	0.6294	0.8061
INSTRE	0.0665	0.2826
Sculpture	0.1620	0.5384

Results



DL Based Model-2

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

DL Pipeline 2

- Partially Implementemented Bags of Local Convolutional Features for Scalable Instance Search and Faster RCNN for Instance retrieval used faster RCNN for global ranking.
- Similarity matching using cosine similarity
- Global and Local Ranking approach as used in DL model-1
- Ranking one using Faster RCNN model and local re-ranking using Bag of Words approach of windowing.
- **Limitations of Bag of Words Model**

  High Dimensional features mostly zeroes

  Lag in Ranking so trade off between time and accuracy

Results

- On ranking using F_CNN convolutional features , it gives a dense representation.
- But evaluation metrics suffers due to dense representation

Faster RCNN Results

Results:

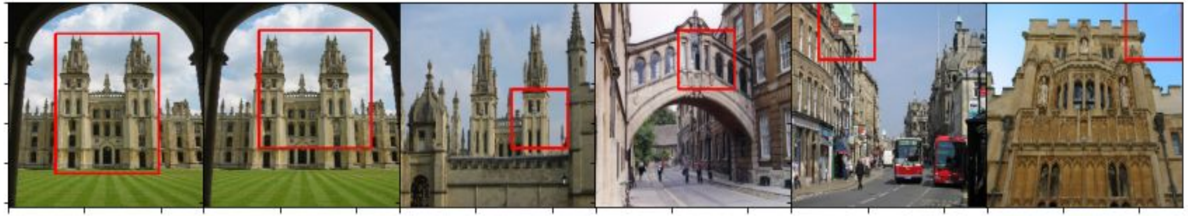
SIMILARITY : COSINE

Dataset	MAP	NDCG
Oxford	0.6515	0.7359
Paris	0.5608	0.7652
INSTRE	0.0909	0.2733
Sculpture	0.1324	0.5829

SIMILARITY : WEIGHTED COSINE

Dataset	MAP	NDCG
Oxford	0.6214	0.7050
Paris	0.5008	0.7251
INSTRE	0.0709	0.2533
Sculpture	0.1314	0.5629

Results:



Analysis

- Bag of Word CNN is performing better than all other model. In terms of evaluation.
- ML features were not good in capturing all details.
- Weighted Cosine giving similar results on evaluation

Future Work

Limitations and future work

We can further use RNN based memory mapping as a future work which can remember features

We can also apply meta learning to model to adopt generalized concepts