Ahmed Alhaj
CMSC 476
Paper project

Image retrieval is one of the most interesting research areas now in the field. Content-based image retrieval (CBIR) systems are used in order to automatically index, search, retrieve and browse image. Colour and texture features are important properties in content-based image retrieval systems.

In content based image retrieval, there are many challenges that have not been fully addressed. Some of the essential challenges, the difficulty that a user suffers to express the expected visual content by a query at hand, such as example image or a sketch map. second , the difficulty in describing high-level-semantics content with low-level visual feature. The time with processing the image query and exact sort of lower-high dimensional representation. Most importantly it extremely hard to develop a unified general content based image retrieval that capable of efficient to all purpose of image retrieval for example a system that can be used to in clinic lab to identify cancer cells meanwhile can identify and be use for surveillance system.

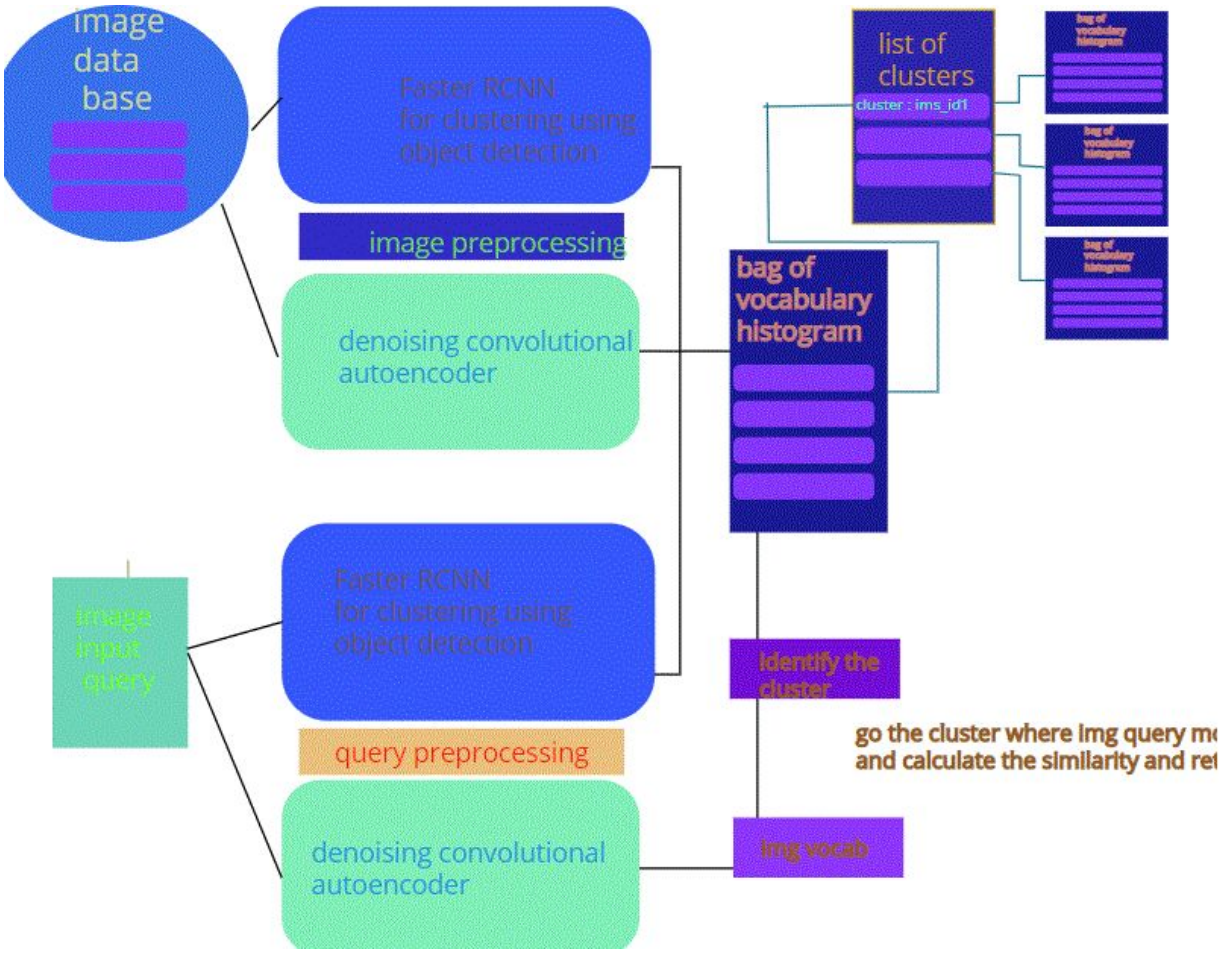
During the last decade, there have been extensive research on content-based image search. The progress in those years has been comprehensively in several areas. The most influential works in the last decades are the development of the SIFT scale invariant features transform and the use of the bag of visual words.

Previous work paradigm

The general cbir architecture for a system goes as following.
The first task in Cbir system is building the features index for the images. The step involves processing and the image and extract feature and storing them in an image inverted index or any equivalent structures. to store In doing so it is important to identify the techniques and the type of feature that needs to be extracted. Previously

people have been relying list of image descriptors most commonly the sift, hog, lbf, surf pca, for feature extractions. These descriptor are mostly evaluated based on four criteria which are the illumination, affine transformation, rotations. So these descriptor have some various performance with respect to these criteria. Then bag of words model is being used to generate vocabulary histogram for each image descriptor. This sort of is being used in different flavor, several research proposed the idea features fusion of different descriptors which is shown to increase precision. the reason for that is that some image-descriptors have more tendency to be resilience with respect to criteria and weak with other criteria. The downside with the idea of descriptor fusion is that feature vector is getting too large and as a result the computation cost for measuring the similarity between 2 vector gets costly. The similarity comparison between feature vector is very challenging problem especially for large multipurpose system, but there are general technique such as euclidean distance and hamming distance are mostly used.

Some of the challenges with this paradigm cost, efficient of retrieval, and scale of system. So as the system gets larger this sort of paradigm gets inefficient and slow and so inaccurate especially with emergence of deep learning and advance of hardware computation speed in gpu.



This method

In this project, I propose simple architecture for resilience cbir model that includes image features representation, database indexing, and image scoring.
The goal do comparison cbir based on fast R-cnn and Denoising Convnet feature extractor. Which will enlist the the development of both model although with R-CNN model which will be Vgg16 keras built in model. And With regard to image image index I will be using simple inverted index like structure where the key is the cluster and the key list is images belongs that cluster. And lastly applying recall and precision method for both models.

Images Representation

This architecture enlists the the generation image vocabulary, development of a set clusters where for each clusters will be a vector of image-vocabulary ID. word vector for each image a label of image content.
This method uses VGG16 fast RCNN models which is modified version of AlexNet and has shown extremely good performance in object detection. And nowadays is consider the state of art in object detection and image classifications. An advance version of fast R-CNN have been which uses a Spatial Pyramid based pooling layer to improve classification and detection performance.

The initial version is known as Fast R-CNN which uses search selective to find out the region of interests and passes and feed them into to ConvNet. It tries find out areas that might be an object by combining similar pixels and textures into several rectangles. And the final output is fed into SVM for classification and then regression between predicted bounding rectangles are computed. For the purpose of this project will be extract feature vector from last layer.

Denoising autoencoder

A denoising autoencoder is a feed forward neural network that learns to denoise images.

By doing so the neural network learns interesting features on the images used to train it. Then it can be used to extract features from similar images to the training set.

Denoising an image. This first step to train the to denoise an input test image First we regenerate the noisy data and load the previously trained autoencoder.

Feature Index

Aside from good features extraction, feature indexing and similarity measurement matrix also play vital roles to facilitate the execution of queries. In general, feature indexing refers to a database organizing structure to assist fast retrieval process. There are two commonly used indexing techniques in CBIR are inverted file index and hashing based indexing.

Relative work

This area of building feature index is researched heavily and the state some of art techniques are partition clustering such as K Means and hierarchical clustering such as agglomerative clustering. the is to build the index based on image clusters with has id,image path, feature vector. A recent proposed techniques such as Joint Unsupervised Learning (JULE), which recursive cnn model
Hierarchically nested data clusters are structured where data clusters at higher layers. represent one or multiple clusters at a lower layer based on mean values of the cluster centers. The first layer clusters are generated based on feature from CNN model.

This index in this method

In this project, the index will built by using partition clustering kmeans. This way is done it cluster the convolutional features vectors and denoising convolutional autoencoder for image sets into two different index. Then i will be trying to fuse both feature vectors

and build index by clustering fused feature vector. The way the index is build is that each cluster the will map to a dictionary key and image feature vector, id, path will be to dictionary value as object holds all image info. So when the query image will be processed using and the convolutional feature vector extracted, and clustered, it will be just look at images in that particular cluster in the index and then compute the similarity score and then sort them. Then it will be just plot first 10 images out of sorted score since they have highest score.

Comparison

For the comparison measures I will be use the scikit learn function label_ranking_average_precision_score..

In this method, I will be computing the relevance score from the computed distance between the feature of the query image and the images of the cluster in the feature index. The lower the distance the higher the relevance score should be.
This scoring function returns a maximum score of 1 if the closest images have the same label as the query image. If there are images with a different label that are closer to the query image, the score decreases. And then report the score with respect to convnet feature index and denoising conv autoencoder feature index.

Project will be done on python
Libraries used will include numpy scipy sklearn skimage kereas and opencv, collections and some other libraries.
Dataset: so far I have label dataset of people faces. done facial image data it was in the spring when I start looking evaluating the project time. So I will be in touch to show you my progress.
https://arxiv.org/pdf/1706.06064v2.pdf

orenstein, t. (2017). test paper. *test journal*.
Wengang Zhou, Houqiang Li, and Qi Tian Fellow, IEEE. (2017). Recent Advance in Content-based Image. 22.

A. Alzubi, A. Amira, and N. Ramzan, “Semantic content-based image retrieval: A comprehensive study,” Journal of Visual Communication and Image Representation, vol. 32, pp. 20–54, 2015

B. Zafar, R. Ashraf, N. Ali, M. Ahmed, S. Jabbar, K. Naseer, et al., "Intelligent Image Classification-Based on Spatial Weighted Histograms of Concentric Circles," Computer Science & Information Systems, vol. 15, 2018.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4912113/
N. Ali, K. B. Bajwa, R. Sablatnig, S. A. Chatzichristofis, Z. Iqbal, M. Rashid, et al., "A novel image retrieval based on visual words integration of SIFT and SURF," PloS one, vol. 11, p. e0157428, 2016.

Scalable Database Indexing and Fast Image Retrieval Based on Deep Learning and Hierarchically Nested Structure Applied to Remote Sensing an
Plant Biology
Pouria Sadeghi-Tehran 1,* , Plamen Angelov 2 , Nicolas Virlet 1 and Malcolm J. Hawkesford