DSAA
Project Summary Report

Group no. : 1
Group members:

1. Riya Jain
2. Aaquib Niaz
3. Alivelu Sreevidya
4. Vijendra Kumar Saini

Project: Content Based Image Retrieval

Background:

Content-based image retrieval (CBIR) is regarded as one of the most effective ways of accessing visual data . Also known as Query By **Image** Content (QBIC),CBIR is the application of computer vision techniques to the image retrieval problem.Content-based image retrieval is opposed to traditional **concept-based approaches.**"Content-based" means that the search analyzes the contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. The term "content" in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself.

Problem Statement:

Given a query image, to perform efficient retrieval of related images from the database and predict the class of the image.

Initial Approach:

Initially, we read various papers and articles about the traditional methods for

Content Based Image Retrieval.As a result, first approach that we worked upon was Support Vector Machines(SVM).

One of the papers we referred to:

http://web.cs.iastate.edu/~honavar/hearst-svm.pdf

Dataset:

The dataset used in this project has been taken from:

Link : http://wang.ist.psu.edu/docs/related/

The dataset consists of images of 10 categories. Each category has 100 images, hence it has 1000 images in total. The categories are : Africa, Beach, Monuments, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains, and Food.

Initial results:

For implementing SVM, we first performed feature extraction of images.

Feature Extraction

We start by mapping the pixel values to the feature space. Hence, feature matrix of the dataset images has been constructed. Each image has 190 dimensional feature vector. Features evaluated for an image are :

- Colour Histogram : HSV space is chosen. Each H,S,V component is quantised into 8,2,2 bins respectively.
  - 32 dimensions
- Color auto-correlogram : The image is quantized into 4x4x4 = 64 colors in the RGB space.
  - 64 dimensions.
- Color moments : The first two moments (mean and standard deviation) from the R, G, B color channels are extracted
    - 6 dimensions

- Gabor wavelet: Gabor wavelet filters spanning four scales: 0.05, 0.1, 0.2, 0.4 and six orientations are applied to the image. The mean and standard deviation of the Gabor wavelet coefficients are used to form the feature vector
    - 48 dimensions
- Wavelet moments : Applying the wavelet transform to the image with a 3-level decomposition, the mean and the standard deviation of the transform coefficients are used to form the feature vector
    - 40 dimensions

Classification:

Initially, we used SVM algorithm for classification. In this approach, **svm(one-vs-one)** has been used for prediction.For that reason n!/ (n-k)! k! binary classifiers have been trained for that particular task.The final matrix has dimensions 1000x192 where the 191st column corresponds to the name of the image file and the 192nd column corresponds to the true class label.

During the training 'hold out' cross validation is performed to increase accuracy where 50% of the samples is used for training and the other 50% for testing.

Challenges faced:

Using SVM for prediction didn't yield a good accuracy for the given dataset. The accuracy we got by applying SVM for classifying images ranged between 50% to60%. Also, the system didn't function properly for query images outside the dataset.

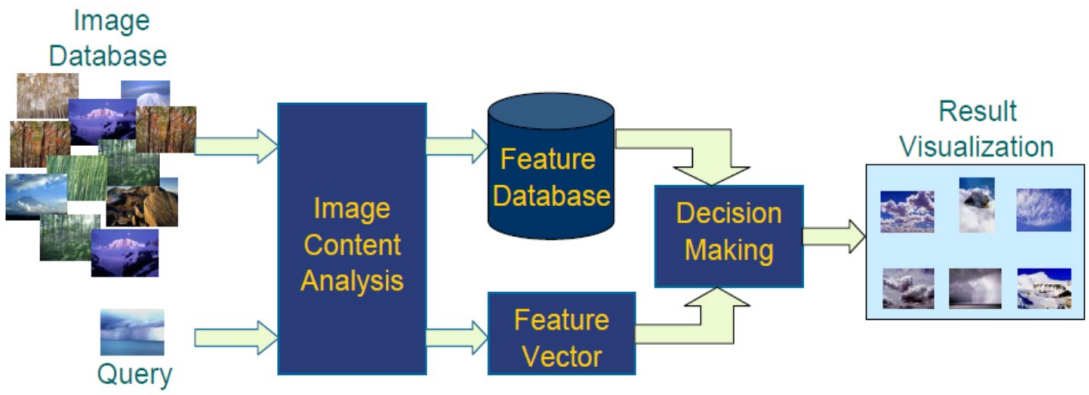
Solutions:

- **Approach to enhance accuracy:**
    1. Since, accuracy using SVM wasn't sufficiently good, we tried applying an SVM+KNN based approach. Here, we first predicted the class of the query using SVM and then applied KNN on that class and query image. This enhanced the accuracy of our system drastically.
    2. Instead of polynomial or gaussian kernel function which we were using earlier, we now used linear kernel function which gave better results.
    3. In addition to using KNN along with SVM, we tried different distance functions to predict the class of a query image. Some of the distance functions we used were: Chebyshev, Cosine, Correlation, Spearman, Normalized L2, Minkowski, Relative Deviation,etc.

       After using these techniques, accuracy improved to about 70%.

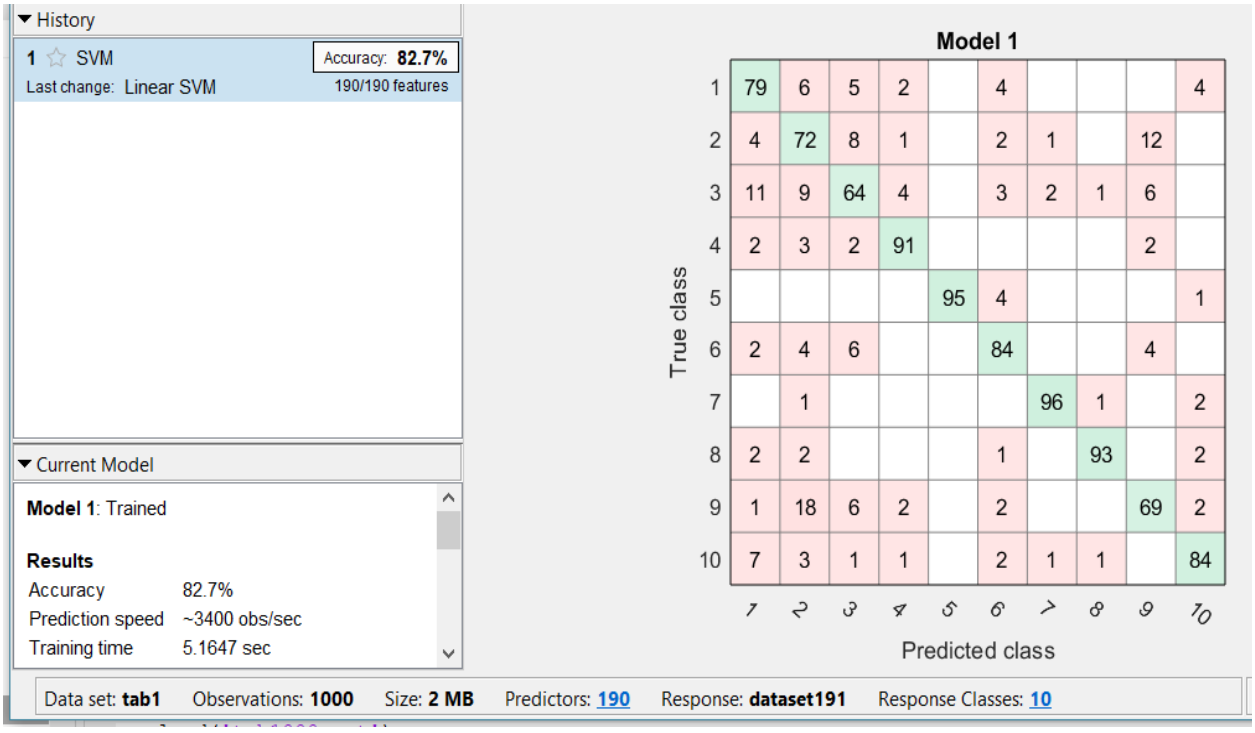
- Bugs were removed so that the system now works for query images outside the trained dataset.

Schematic Overview of the system:



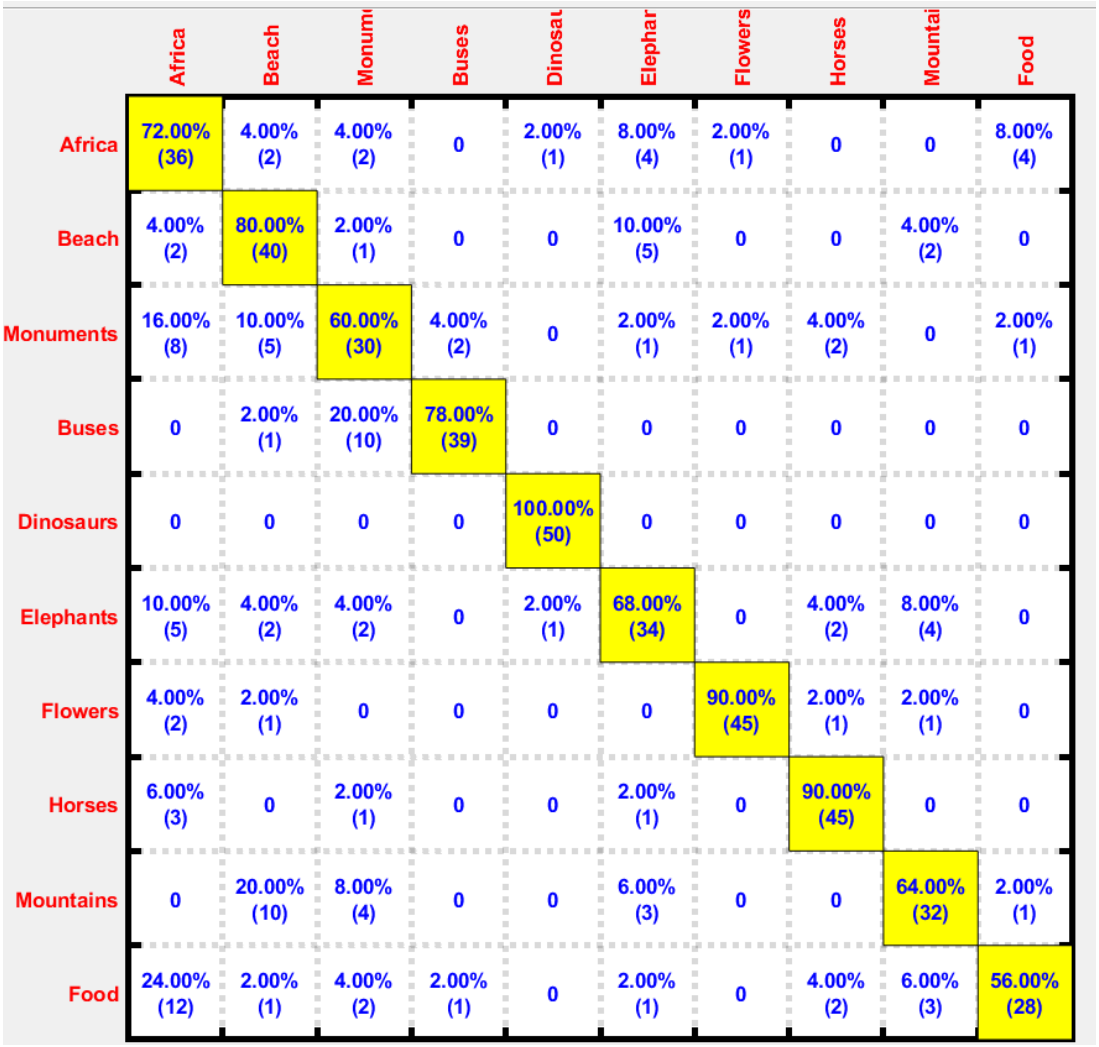
Results:

Accuracy of the training model using MATLAB's classification learner : **82.7%.** On the right hand side is the confusion matrix.



Confusion Matrix of the 'hold out' cross validation carried out during

training. The dataset is split into two equal halves for this cross validation.



	Africa	Beach	Monumi	Buses	Dinosaı	Elephar	Flowers	Horses	Mountai	Food
Africa	72.00% (36)	4.00% (2)	4.00% (2)	0	2.00% (1)	8.00% (4)	2.00% (1)	0	0	8.00% (4)
Beach	4.00% (2)	80.00% (40)	2.00% (1)	0	0	10.00% (5)	0	0	4.00% (2)	0
Monuments	16.00% (8)	10.00% (5)	60.00% (30)	4.00% (2)	0	2.00% (1)	2.00% (1)	4.00% (2)	0	2.00% (1)
Buses	0	2.00% (1)	20.00% (10)	78.00% (39)	0	0	0	0	0	0
Dinosaurs	0	0	0	0	100.00% (50)	0	0	0	0	0
Elephants	10.00% (5)	4.00% (2)	4.00% (2)	0	2.00% (1)	68.00% (34)	0	4.00% (2)	8.00% (4)	0
Flowers	4.00% (2)	2.00% (1)	0	0	0	0	90.00% (45)	2.00% (1)	2.00% (1)	0
Horses	6.00% (3)	0	2.00% (1)	0	0	2.00% (1)	0	90.00% (45)	0	0
Mountains	0	20.00% (10)	8.00% (4)	0	0	6.00% (3)	0	0	64.00% (32)	2.00% (1)
Food	24.00% (12)	2.00% (1)	4.00% (2)	2.00% (1)	0	2.00% (1)	0	4.00% (2)	6.00% (3)	56.00% (28)

Analysis and Findings:

Observations showed that the accuracy of the system was determined by the following factors:

- **Algorithm:** We have tried and implemented 3 approaches: SVM , and KNN, and first SVM and then KNN(on the class predicted by SVM) for classification, we came to the following conclusion:
    - The knn method is faster, and has a better ability to adapt its model to previously unseen data. However, knn fails to produce

good results with noisy, irrelevant, nominal or missing attribute values. Moreover, the hypothesis complexity can grow with the data. In contrast, using svm and then knn for the class predicted by svm gives a better accuracy over the knn method.

- **Kernel Function:** After trying polynomial, linear, and gaussian kernel function in fitcsvm, we found out that 'linear' kernel function gave best results.
- **Distance Function:** It has a major impact on the accuracy of the system. E.g. Cosine, Chi squared distance functions have better accuracy over other distance functions.

Challenges still unresolved:

The accuracy of the cbir is now 70%. However, though the system now works for query images outside the dataset, the accuracy still lies between 40 to 45 %.