BITS G540: Research Practice

Second Semester 2022-23

Project Report

Research Topic: Content Based Image Retrieval

Instructor in charge - Dr. Avinash Gautam

SUBMITTED BY

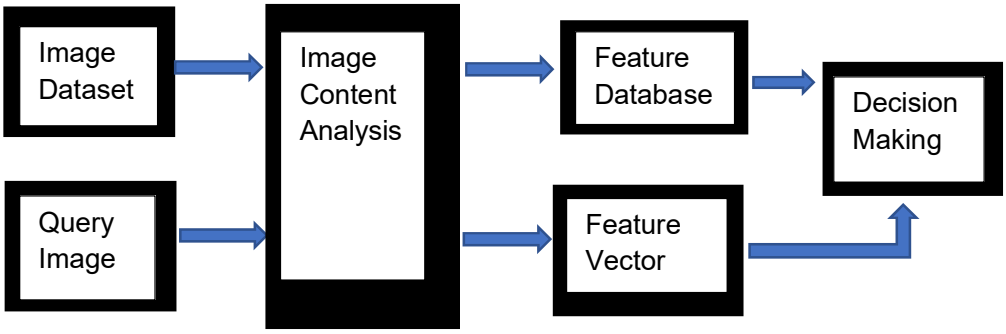
JOSHI KASHYAP - 2022H1030116P

Contents: -

- Introduction
- Image Data Set
- Feature Extraction
- Query by Sample and Classification
- Results
- Observations
- Conclusion & References

1.Introduction: -

The purpose of this project is to develop an efficient image retrieval system that utilizes data preprocessing techniques and classification algorithms. With the increasing amount of multimedia data available on the internet, there has been a growing need for efficient media retrieval systems. In this project, a dataset of images will undergo data preprocessing to convert them into feature vector using feature extraction techniques. The system will then take in a query image and use K-nearest neighbors (KNN) and support vector machine (SVM) classification algorithms to return similar images and predict their labels. The accuracy of the system will be evaluated by comparing the predicted labels with the actual labels for testing purposes.



[Flow Diagram Of Project]

2.Image Dataset: -

The image dataset used in this project consists of two sets of images. The first dataset contains 1000 images from 10 different categories, with 100 images corresponding to each category. The categories are buses, beaches, dinosaurs, elephants, flowers, food, etc. Each image is in JPEG, JPG or in PNG format and has a size 2 KB to 40KB.

Data Set No.1
Categories

Human	Beach	Bus	Dinosaur	Elephant	Flower	Food	Horse	Monument	Mountain

The second dataset contains 1200 images of different celebrity faces, including George Bush, Salman Khan, Virat Kohli, and others. The images were obtained from publicly available sources and have a 2 KB to 30KB. Each image is in JPEG, JPG or in PNG format.

Data Set No.2
Faces Data Set
Categories

Virat Kohli	George Bush	Tom Daschle	Salman Khan	Tom Cruise	Priyanka Chopra	Michelle Obama	Maria Kondo	Jackie Chan	Barack Obama

Both datasets were selected to represent different image types and to demonstrate the applicability of the image retrieval system across a range of domains. The images were manually labeled with their corresponding category or celebrity name to facilitate the classification and retrieval process.

Prior to preprocessing, the datasets were randomly divided into training and testing sets, with 80% of the images used for training and 20% for testing. This partitioning was used to evaluate the accuracy of the image retrieval system and to ensure that the testing images were not used during the training phase.

3.Feature Extraction

After obtaining the image dataset, the next step in the project is feature extraction. The goal of feature extraction is to transform each image from its raw pixel format into a format that can be understood by the prediction algorithms. This process involves a set of six steps that are applied to each image in the dataset.

The first two steps involve computing the color histogram and color auto-correlogram for each image. The color histogram captures the distribution of colors in the image, while the color auto-correlogram captures the spatial correlation of colors. These features provide important information about the color content of the images. The third step involves extracting the mean and standard deviation of each RGB color channel. This step helps capture the basic color statistics of the image.

Moving on to the fourth step, the mean and standard deviation of the Gabor wavelet coefficients are computed for each image. Gabor wavelets are known for their ability to capture the texture and frequency content of images.

The fifth step involves applying the wavelet transform to each image with a 3-level decomposition, and extracting the mean and standard deviation of the transform coefficients. This step helps capture the texture and frequency content of the image at multiple scales.

Features	**Description**	**Dimensions**
Color Histogram	HSV space is chosen, each H, S, V component is uniformly quantized into 8, 2 and 2 bins respectively	32
Color auto-correlogram	The image is quantized into 4x4x4 = 64 colors in the RGB space	64
Color moments	The first two moments (mean and standard deviation) from the R, G, B color channels are extracted	6
Gabor wavelet	Gabor wavelet filters spanning four scales: 0.05, 0.1, 0.2, 0.4 and six orientations: $\theta 0 = 0$, $\theta n+1 = \theta n + 6\ \pi$ are applied to the image. The mean and	48

	standard deviation of the Gabor wavelet coefficients are used to form the feature vector	
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

Finally, in the sixth step, all the feature vectors from the previous steps are combined into a new concatenated vector that is used to represent each image. This concatenated vector has a dimensionality of 32+64+6+48+40+1 and is utilized in the subsequent steps of the project to perform image classification and similarity retrieval.

4.Query by Sample and Classification

In research we have included two use cases. First use case is to find *n* similar images and second use case is classification of images and measure the accuracy. For this use cases K-Nearest Neighbors (KNN) algorithm and Support Vector Machine (SVM) algorithm is chosen.

So, the user is able to insert an image into the code and retrieve similar images from the dataset, as well as predict the label of the image using K-Nearest Neighbors (KNN) algorithm. The KNN algorithm is a type of classification algorithm that is based on the idea of finding the k nearest neighbors of a sample in a feature space and assigning it the label that is most common among those neighbors.

Additionally, the dataset is divided into training and testing sets, and Support Vector Machine (SVM) algorithm is utilized for classification. The SVM algorithm is a type of supervised learning algorithm that is used for classification, regression and outlier detection. It works by finding a hyperplane that separates the data into different classes with the maximum margin between them.

The trained SVM algorithm is used to predict the label of the query image and the testing part of the dataset is run to calculate the accuracy of the classification. This step provides a way to retrieve similar images and predict the label of a query image, allowing for efficient classification and retrieval of images.

5.Results

It was observed that the accuracy achieved for the first dataset, which contains around 1000 images, is **80.5%**. This accuracy was obtained by applying the query by sample and classification approach using the KNN and SVM algorithms. The KNN algorithm was able to retrieve similar images based on the feature vectors obtained from the image dataset. On the other hand, the SVM algorithm was able to classify the images and predict their labels with a good level of accuracy.

In the second dataset, which contains around 1200 images of faces, the accuracy achieved was around **87%**. This accuracy was obtained by applying the same approach as in the first dataset.

6.Observations

- Currently using a dataset that contains 1000 and 1200 images, which is not sufficiently large enough to get accurate results.
- Features extraction techniques which we are currently using are efficient, but higher resolution images need more features to be included for more accurate results.
- For better results in faces dataset several additional feature extraction techniques like Local Binary Patterns can be introduced.
- One major drawback of KNN is its sensitivity to high-dimensional data, such as images. where the distances between data points become large, and their relative ordering becomes meaningless.
- KNN also requires a lot of computational resources to calculate distances between each pair of data points in the training set, which can be prohibitive in large datasets.
- SVM algorithm can be computationally expensive, especially when dealing with large datasets and high-dimensional feature spaces. In order to perform multi-class classification, multiple SVM classifiers must be trained, which can further increase computational complexity.
- SVM can be sensitive to class imbalance, where one class has significantly fewer samples than another, leading to biased results.

7. Conclusion & References

In conclusion, this project successfully implemented a system for image classification and similarity retrieval using various feature extraction techniques and machine learning algorithms. The results showed that the system achieved high accuracy in classifying images and retrieving similar ones. However, there were limitations with the dataset size and computational resources required by some of the algorithms used. Overall, this project demonstrates the effectiveness of feature extraction and machine learning techniques for image analysis tasks.

Reference paper: https://arxiv.org/pdf/1608.03811.pdf
Image Data Set No.1:ImageDataSet1_Structures
Image Data Set No.2: ImageDataSet2_Faces