DEVELOPING DYNAMIC ANNOTATION USING BACK PROPAGATION CLASSIFICATION TECHNIQUE IN CONTENT-BASED IMAGE RETRIEVAL

Saksham Rai

ABSTRACT

The CBIR will in general search, retrieve and index an image. The positioning of a CBIR framework relies upon the impossible to miss image representation, design, highlights and so on. CBIR keeps away from numerous issues related to conventional strategies for recovering images by utilizing keywords. The explanation of a CBIR framework mainly depends on the specific image portrayal and coordinating closeness capacities utilized [1]. Content-Based Image Retrieval framework depends on the recovery of the indistinguishable pictures from the wide database when a query picture is a feed as a contribution to the framework. The entire working of Content-Based Image Retrieval (CBIR) framework is to give comparable matches of a question contribution from enormous and various datasets. Profitable sanctioning of a CBIR framework requires arrangement, characterization, ordering, and repossession of images.

***Keywords**: CBIR, retrieval, feature extraction, classifier, Back Propagation Neural Network (BPNN)*

I. INTRODUCTION

Being an easy, effective and less costly image extraction method, CBIR has become interesting as well as challenging problem with numerous applications with the availability of huge video and image data sets. Term Content in CBIR may allude to shape, shading, surface, direction, or some other data got from the picture [2]. The main discipline of depicting color information of images in CBIR systems is by use of color histogram [3]. A Colour histogram is a type of bar graph, in which each bar represents a specific color being used uploaded image or the image to be tested. The bars in a color histogram are termed to be bins, and they are represented by the x-axis. The number of bins leans on the number of colors present in an image. The number of pixels is indicated by the y-axis. CBIR system gives top matching output to query image from a huge database [4]. Effective CBIR system categorizes, index and retrieve images, which depends on effective features to describe the content of an image. The large contrast within a class requires alternate features and methods to maximize or improve the retrieval performance of the system.

1.1 Retrieval of image

(1) Colour based retrieval- This feature is the most used feature, and generally histograms are used todescribe it. This method has advantages of speediness, low memory space requirement.

(2) Texture based retrieval- When there is a description of the texture of the image, usually texture's statistic feature and structure features are adopted.

(3) Shape feature retrieval- Shape is related to the uniqueness and description of the object in the image, so shape's semantic feature is much stronger than texture.

1.2 Feature Extraction (SIFT Algorithm)

This component was proposed by David Lowe [5] that has the ability to separate and depict neighborhood components of picture effectively. The SIFT incorporates five phases. The first stage searches for extreme points of the image. Localization of key points is done at the second stage to figure out the exact shape of the image. Orientation around each key point is done with the help of histograms. Estimated key points describe the gradient and magnitude of image. Finally trimming of erroneous coordinates is done.

Image classification is the basis of learning in which attributes are matched to the feature vector space, as shown in figure 1.1 [6]. These algorithms are called machine learning algorithms.

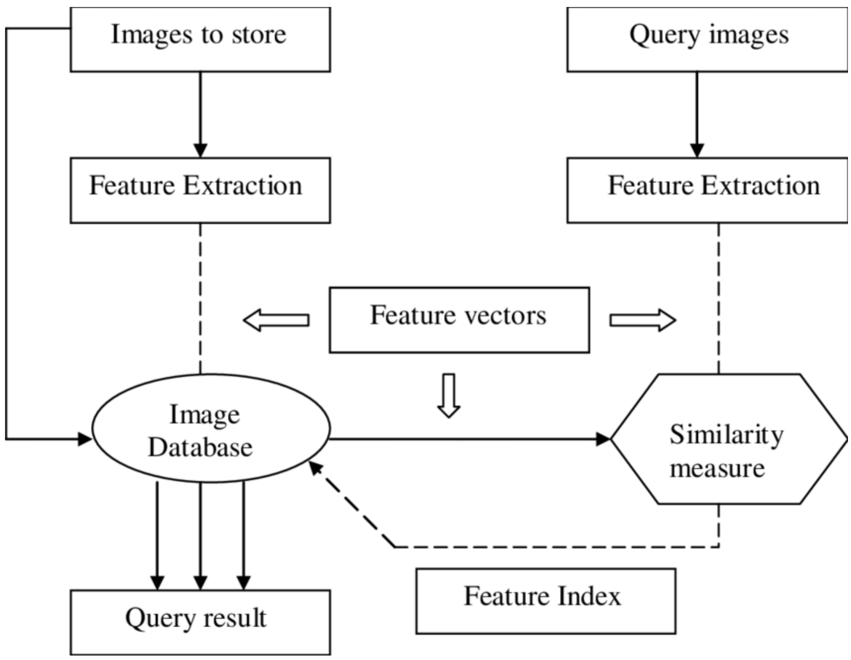

Fig 1: Classification Technique

Two stages in classification are as follows:

(1) Training stage- In the training section, classifier learns classification rules by its own.

(2) Testing stage- In the testing phase, the feature vectors of a query input are taken as input.

1.3 Use of Multi-layer Back Propagation Neural Network

BPNN is an artificial neural network technique based on the error backpropagation algorithm. The Back Propagation, Neural Network model, consists of three layers: an input layer, some hidden layers & an output layer. A neuron at every layer has a distinctive weight. Input neurons obtain input information from exterior sources and then go to a hidden layer which is an interior info processing layer & is answerable [7] for the information conversion, and then the nodes in the production layer supply the required output substantially. Every weight is reviewed and backpropagated layer by layer from the output layer to hidden layer & input layer. This procedure will be sustained until the output error of the network is reduced to a tolerable level.

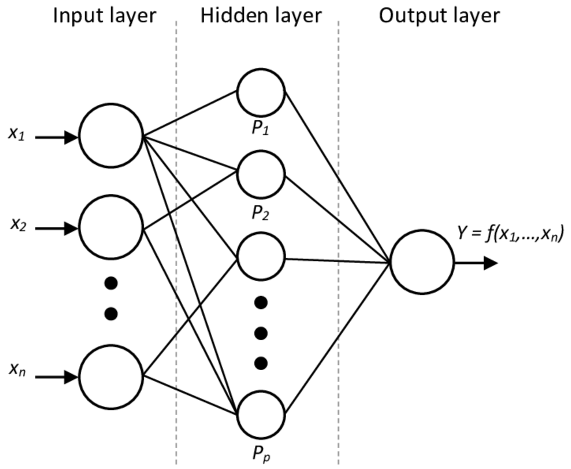

Fig 2: Multilayer

PROBLEM FORMULATION

There is a semantic gap in result image due to lack of coincidence between the data that one can excerpt from the visual information and the understanding that the same data have for the user in a given situation. Sometimes there are several kinds of variations in a picture that affect the classification. The semantic gap between the client requirement and the capability of CBIR algorithms remains significant. Significant efforts have been put into low-level image properties such as color and texture. To handle the problem of semantic gap faced by CBIR, learning semantics from training data input and developing retrieval mechanisms to leverage semantic estimation are important directions correctly. Most methods use them as part of histograms over the whole segmented image. Lighting reasons important variations in the strength of the pixels. Lighting alteration in the picture has a key ascendancy on its aspect. Illumination and the existence of shadows is also a problem in CBIR. Several numbers of times altered the appearance of objects which makes the recognition of image difficult for the system to compare it with the test image for accurate results. The major drawback of CBIR approach is that do not heed another lesson from Pattern Noisy background may hinder the performance. This problematic increases false-positive results in result image retrieval. Image retrieval is very hard as compared to text retrieval. The reason for such difficulty is that the image data (pixels stored in bytes) are very far from the interpretation of images by human observers.

OBJECTIVES

The set objectives of the set to be achieved in the proposed model are as following:

1. To study the various CBIR system development techniques.

2. To extract the features of the database [13] using the SIFT algorithm and to train the system using the extracted feature set, also to set the target of the extracted features for the Multi-Layer Neural Network.

3. To evaluate the test image over the targeted dataset (using BPNN).

4. To evaluate the performance parameters like FAR, FRR, Accuracy, Mean Square Error, Precision, and Recall.

5. To compare the accuracy of the proposed model with previously implemented models.

METHODOLOGY

Dynamic Annotation system for Content-Based Image Retrieval is employed using SIFT and Back Propagation Neural Network. Various CBIR techniques study has been carried out. Datasets of various images are created at initial and runtime using UCI learning machine. Each data set is loaded with numerous images of a different type. Various categories of datasets are represented in Graphical User Interface. Afterimage upload, the histogram is generated to find the maximum intensity of an image. After this step, conversion of the original image is done into a greyscale image in order to reduce the number of pixels. Using Scale Invariant Feature Transform, features of all datasets are extracted in order to train the system using extracted features of datasets. SIFT algorithm finds all the key points and identifies unique image properties.

Basic step SIFT algorithm is to to build a Gaussian scale space, which includes an adaptable scale 2-Dimensional Gaussian operator $G(a1, b1, \sigma)$ with the input picture.

$J(b1, c1)$: $M(b1, c1, \sigma) = H(b1, c1, \sigma) * J(b1, c1)$

A difference of Gaussian images $E(b1, c1, \sigma)$ is then obtained by subtracting subsequent scales in each octave:

$E(b1, b1, \sigma) = M(b1, c1, k\sigma) - M(b1, c1, \sigma)$ Where k is a constant multiplicative factor in scale space. Nearby extreme points are then identified by examining each image point in $E(b; c; \sigma)$. A point is a local lowest or extreme when its value is smaller or larger than all its nearby adjoining points.

The target is set to the extracted features for the Neural Network. The test image is uploaded to the system to evaluate it over targeted dataset using BPNN. Different performance parameters are evaluated like False Acceptance Rate, False Acceptance Rate, Accuracy, Mean Square Error, Precision,Recall, and comparisons are done between results of proposed and existing parameters.

EXPERIMENTAL RESULTS

In this section, results are described below:

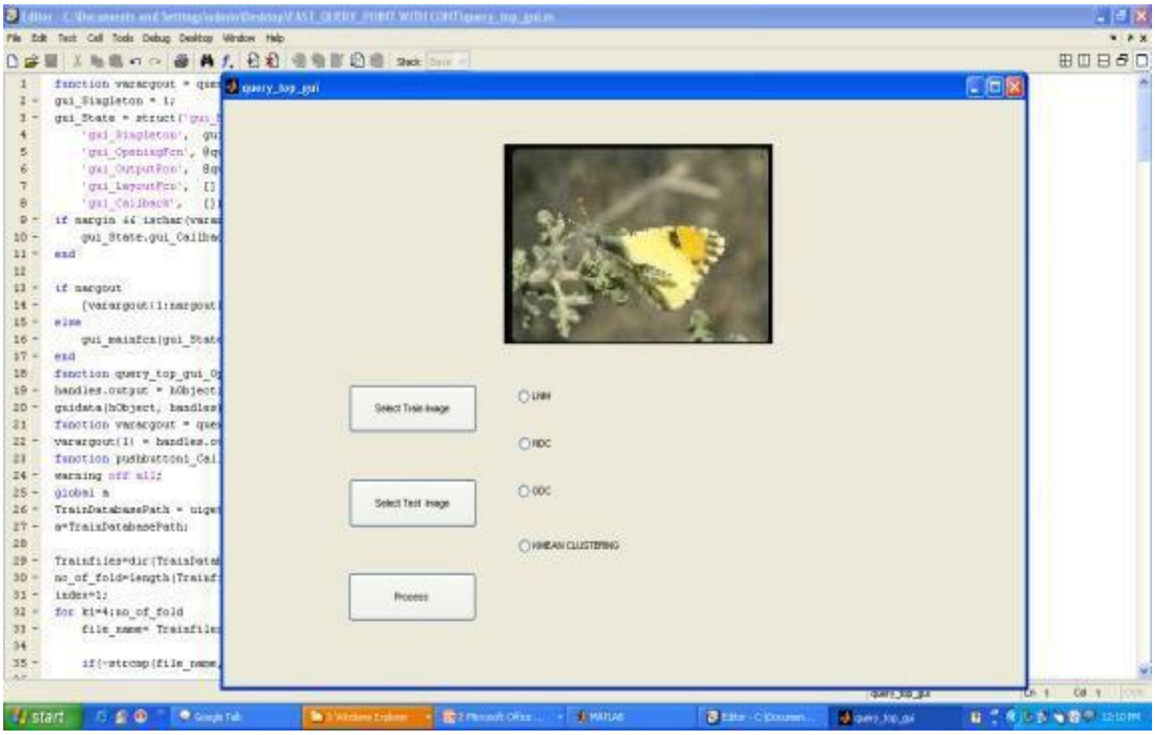

Fig 3 Query Image selection

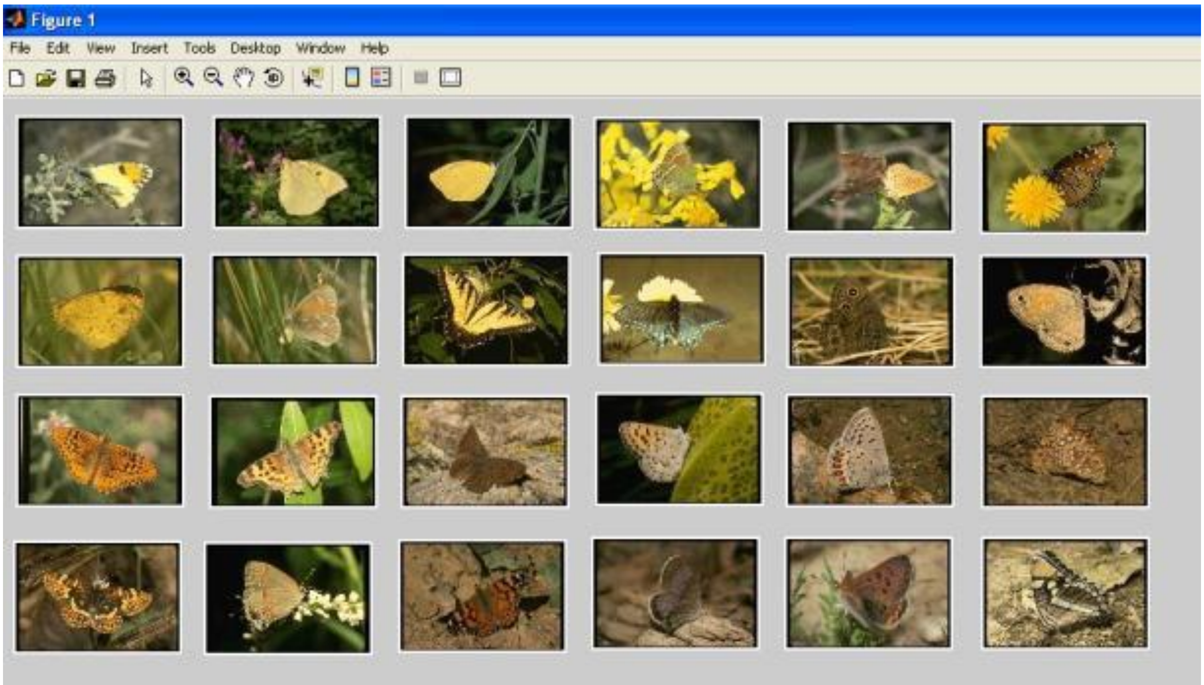

Fig 4:Images retrieved

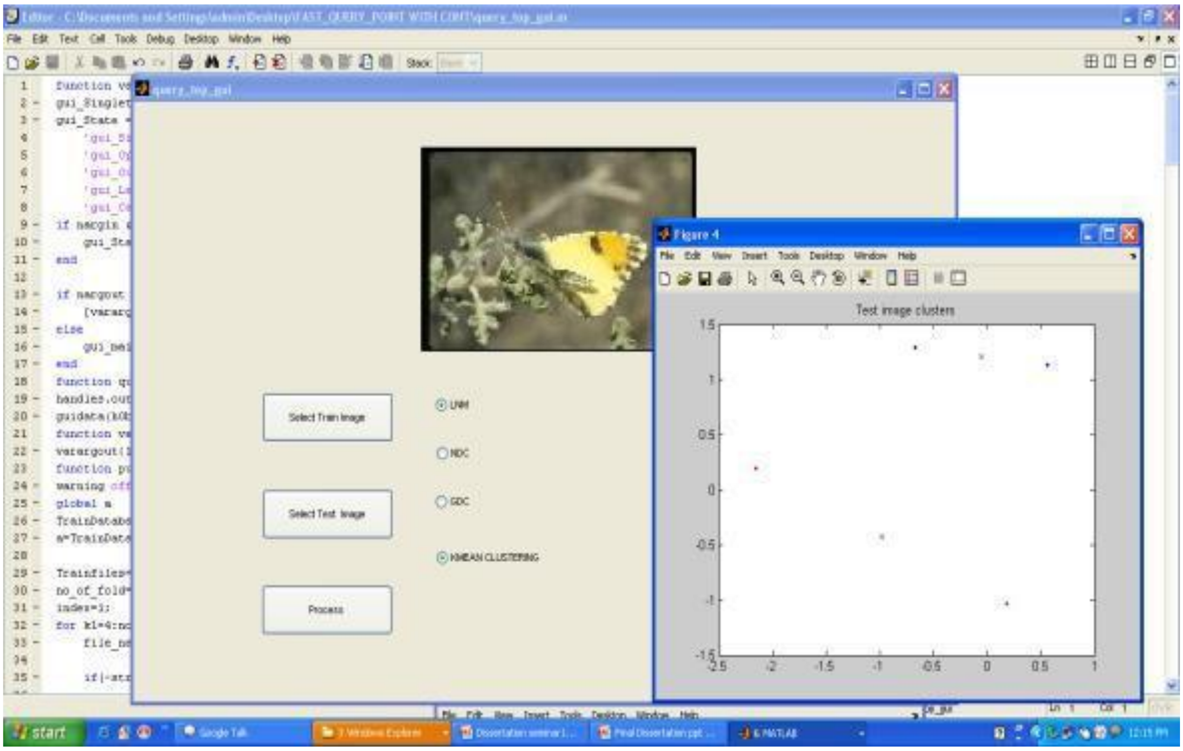

Fig 5:Clustering

As shown in Figure 3, an image from one dataset is uploaded as an input image. The histogram is generated to calculate the intensity of pixels of the image. The input image is converted into a greyscale image after which its key points and edges are detected using SIFT algorithm as shown below in figure 5. The descriptor is recorded at each key- point. The picture gradient, magnitudes, and overviews, with respect to the significant overview of the key factors, are computed.

Figure 5 calculates the minimum value of the precision and recall in existing work. Recall and Precision performance has improved. Maximum precision value performance achieved is 0.98, and the Recall value achieved is 0.99.

CONCLUSION

Motivated by the impediments of such a methodology, ongoing examination inclines in CBIR has continued to interactional frameworks and human PC intuitive interfaces that include a human as a feature of the yield recovery process. In this theory, the issue of picture recovery is illuminated, utilizing a blend of SIFT and multilayer BPNN. Right off the bat, including extraction will be finished utilizing SIFT. Neural system preparing depends on the highlights of pictures in the dataset. The picture highlights considered here are normal worth, least worth, and greatest worth. The preparation is completed utilizing Multi-Layer Back Propagation Neural Network calculation. This prepared framework, when given an inquiry picture recovers and shows the pictures which are pertinent and like a question from the different databases. The outcomes demonstrate an extensive improvement as far as FRR, FAR, Accuracy, Precision, and Recall of pictures recovered.