International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
International Conference and Workshop on Communication, Computing and Virtualization (ICWCCV 2015) – www.ijais.org

# Cloud based CBIR Software as a Service

Krunali Vikas Vartak
Information Technology Department,
Thakur College of Engineering & Technology,
Mumbai, India

Vinayak Ashok Bharadi
Information Technology Department,
Thakur College of Engineering & Technology,
Mumbai, India

## ABSTRACT
The images are used as a prime factor of communication on an effective scale. Hence the need of efficient and effective tools for retrieval of query images from database is increased significantly. CBIR is a technique for retrieving images on the basis of automatically-derived features such as color, texture and shape. Feature extraction used is a technique to extract feature vectors of an image based on color, shape, texture etc. which is generally known as image data. In this paper, cloud based CBIR SaaS architecture is proposed due to which the services of CBIR will be dynamically made available throughout the desired systems resulting in increase in applications scalability, flexibility and availability.

## Keywords
Content Based Image Retrieval (CBIR), Web Role, Blob Storage.

## 1. INTRODUCTION
There are various processes on web through which one can create, process and store images or any other dimensional information. The outcome has resulted into various studies in searching and managing images. Therefore, finding efficient image retrieval mechanisms from large resources has become a wide area of interest to researchers [1]. Image retrieval method is a technique for searching and retrieving images from a large database of digital images. The need to find the desired image from the image database systems which can be geographical maps, pictures, medical images, pictures in medical atlases, pictures obtaining by cameras, microscopes, telescopes, video cameras, paintings, drawings and architectures plans, drawings of industrial parts, space images, etc. is that it is shared by many professional groups, including journalists, design engineers and art historians [2].

Basically the tricks on images are done with the help of an application (software as a service) locally as well as centrally, by requesting the images stored on the image database. The applications deals with the image storing, image retrieving, image processing and they are: Content-Based Visual Information Retrieval (CBVIR) and Content-Based Image Retrieval (CBIR). Content-based retrieval is an important alternative and it is advantageous over traditional keyword-based searching for multimedia such as image and image data. CBIR can greatly enhance pictorial information management as well as supporting with possible opportunities in the form of statistical and comparative analysis of functional image and image data.

CBIR is a technique for retrieving images on the basis of automatically-derived features such as color, texture and shape. Here a query image will be triggered and will be compared to the images stored in the image database. Once the match is found the results are displayed in the form of image output. The CBIR application extracts the image from the image database with the help of input image or image data. The desired image extracted can be used for various purposes such as communication, analysis, guidelines, etc. CBIR application also characterizes image queries into three levels of abstraction: primitive features such as color or shape, logical features such as the identity of objects shown.

## 2. LITERATURE SURVEY
### 2.1 NIR: Content Based Image Retrieval on Cloud Computing
Zhuo YANG, Sei-ichiro KAMATA and Alireza AHRARY in 2009 proposed NIR - an open source cloud based content based image retrieval system. CBIR is one of the challenging and emerging technologies as high computation task because of the algorithm computation complexity and big amount of data. As based on cloud computing infrastructure, NIR is easy to extent and flexible for deployment. Due to this there has been significant increase in scalability and availability. As an open source project, NIR can be improved on demand and integrated to other existing systems [3].

### 2.2 Content-based Image Retrieval (CBIR) using Hybrid Technique
This method of CBIR was researched by Zainab Ibrahim Abood, Israa Jameel Muhsin, and Nabeel Jameel Tawfiq in year 2013. The research delivered Content Based Image Retrieval (CBIR) using four feature extraction techniques. The four techniques used are colored histogram features technique, properties features technique, gray level co-occurrence matrix (GLCM) statistical features technique and hybrid technique. The features are extracted from the database images and query images in order to find the similarity measure between them. The similarity-based matching is post extraction stage in CBIR. The three types of similarity measure used are, normalized Mahalanobis distance, Euclidean distance and Manhattan distance. The research concluded that CBIR using hybrid technique have higher match performance in all kind of similarity measures used [4].

### 2.3 Content Based Image Retrieval Using Fusion of Gabor Magnitude and Modified Block Truncation Coding
Dr. H B Kekre, V A. Bharadi have introduced Content Based Image Retrieval using Fusion of Gabor Magnitude and Modified Block Truncation Coding. Gabor filters comprises of wavelets, where each wavelet captures energy at a specific frequency and a specific direction. Expanding a signal using this basis provides a localized frequency description, therefore capturing local features in the form of energy of the signal.

Texture features can be extracted from the group of energy distributions. And modified block truncation is used to retrieve color feature from image. The proposed system was giving higher Precision and Recall as compared to only Gabor and only MBTC based CBIR. Gabor feature gives good response to texture of the image and Modified BTC give good response to color content of image [5].

### 2.4 Efficient Relevance Feedback for Content-Based Image Retrieval by Mining User Navigation Patterns

Ja-Hwung Su, Wei-Jyun Huang, Philip S. Yu and Vincent S. Tseng in 2011 proposed a method, Navigation-Pattern-based Relevance Feedback (NPRF), to achieve the high efficiency and effectiveness of CBIR in coping with the large-scale image data. In terms of efficiency, the iterations of feedback are reduced substantially by using the navigation patterns discovered from the user query log. In terms of effectiveness, the proposed search algorithm NPRF Search makes use of the discovered navigation patterns and three kinds of query refinement strategies, Query Point Movement (QPM), Query Reweighting (QR), and Query Expansion (QEX), to converge the search space toward the user’s intention effectively. By using NPRF method, high quality of image retrieval on RF can be achieved in a small number of feedbacks. The experimental results reveal that NPRF outperforms other existing methods significantly in terms of precision, coverage, and number of feedbacks [6].

### 2.5 Content Based Image Retrieval Using Independent Component Analysis

Arti Khaparde, B L Deekshatulu, M.Madhavilatha, Zakira Farheen, Sandhya Kumari presented a new approach for global feature extraction using an emerging technique known as Independent Component Analysis (ICA). A comparative study was delivered between ICA feature vectors and Gabor feature vectors for 180 different texture and natural images in a databank. Result analysis show that extracting color and texture information by ICA provides significantly improved results in terms of retrieval accuracy, computational complexity and storage space of feature vectors as compared to Gabor approaches [7].

## 3. WINDOWS AZURE CLOUD AND STORAGE SERVICES

### 3.1 Windows Azure Cloud Services

When creating an application and run it in Azure, the code and configuration together are called an Azure cloud service. A Multi-tier web application can be deployed in Azure by creating a cloud service, defining multiple roles to distribute processing and allowing flexible scaling of the application. A cloud service consists of one or more web roles and/or worker roles, each with its own application files and configuration.

"Web Role" virtual machines are Windows Servers with IIS installed, whereas "Worker Role" virtual machines are Windows Servers without IIS installed. Web roles provide a dedicated Internet Information Services (IIS) web server that can be used for hosting the web front-end of the cloud service. Application code hosted within worker roles can run tasks in the background that are asynchronous, long-running, or perpetual [11].

### 3.2 Windows Azure Storage Services

Windows Azure provides multiple storage services that are highly durable, scalable as well as constantly available. Azure storage provides users with following capabilities to persist both structured as well as unstructured data: Anywhere and anytime access, Store data for any length of time, Scale to store any amount of data, Pay for only what is used/stored. Azure offers three types of storage services, which cater to unstructured, structured as well as transient data requirements, such as: Blob, Table and Queue [12].

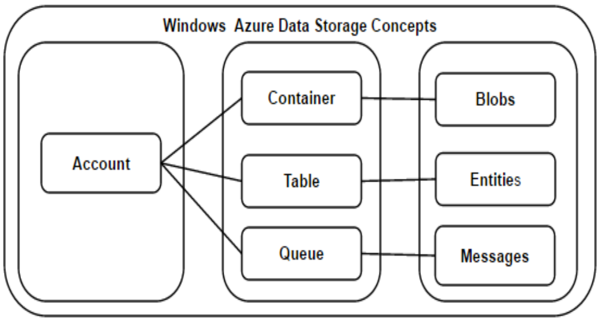


**Fig 1: Windows Azure Storage Services [12]**

Azure Blob storage is a service for storing large amounts of unstructured data, such as text or binary data, that can be accessed from anywhere in the world via HTTP or HTTPS. There are two types of blobs: block blobs and page blob. A “bock blob” can store up to 200 GB of data and is optimized for streaming workloads. While a “page blob” supports up to 1 TB of data and is meant for random access. Each object is stored in a Container and can also have properties associated with it. The properties can be used to store any meta-data information relevant to the object persisted [13].

Queues provide reliable storage and delivery of messages in an application. It helps in storing messages that can be accessed by a client. A queue’s prime function is to enable communication between Web and Worker Role instances. Queues are accessible using REST interfaces. The message size for Azure Queues is limited to less than 8KB [12].

Table storage provides non-relational, schema-less but structured storage facility [12]. It is built to provide massively scalable, constantly available and durable structured storage. In the table storage, a unit of data is persisted as an Entity. Each entity is defined as a set of attributes that constitute the properties of the entity. There is no limit on how many entities can be stored in one table but typically billions of entities can be stored in one table. Also, the storage has no hard limits and should easily scale into the terabytes range. Additionally, entities can also be grouped into partitions. A partition key along with row key are mandatory attributes of any entity object stored in Tables. Partition key is user define value used to logically group related set of entities and a Row key is a user defined unique value within the partition that the entity belongs to. Partitions are required to support scale-out as well as highly durable storage capabilities to the cloud enabled applications on Azure.

## 4. IMPLEMENTATION

CBIR application allows extracting the correct images according to objective visual contents of the image [4][8]. The main aim of the CBIR systems is to provide means to match and find images in large repositories using its contents as low level descriptors. These descriptors do not exactly match the high level semantics of the image; therefore, assessing similarity between two images using only their features is not a trivial task [4][9]. The architecture of CBIR is shown in figure (2).

In this application the request is made in the form of query image. The query image is checked for its size and then its normalized i.e. Image Normalization takes place where image of size m*n is reduced to size 256*256 in order to extract the feature vectors of the image on the basis of color, shape, texture, etc. using image extraction techniques. This technique is implemented by a web service deployed on a Web Role on cloud using Hybrid Wavelet Type1 and Type2. The extracted feature vectors of the query image will be compared with the extracted features of the image residing in the image repository with the help of similarity measures. Similarity measures like precision and recall are used for performance evaluation comparison.

$$\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}} \qquad [10]$$

$$\text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images in Database}} \qquad [10]$$

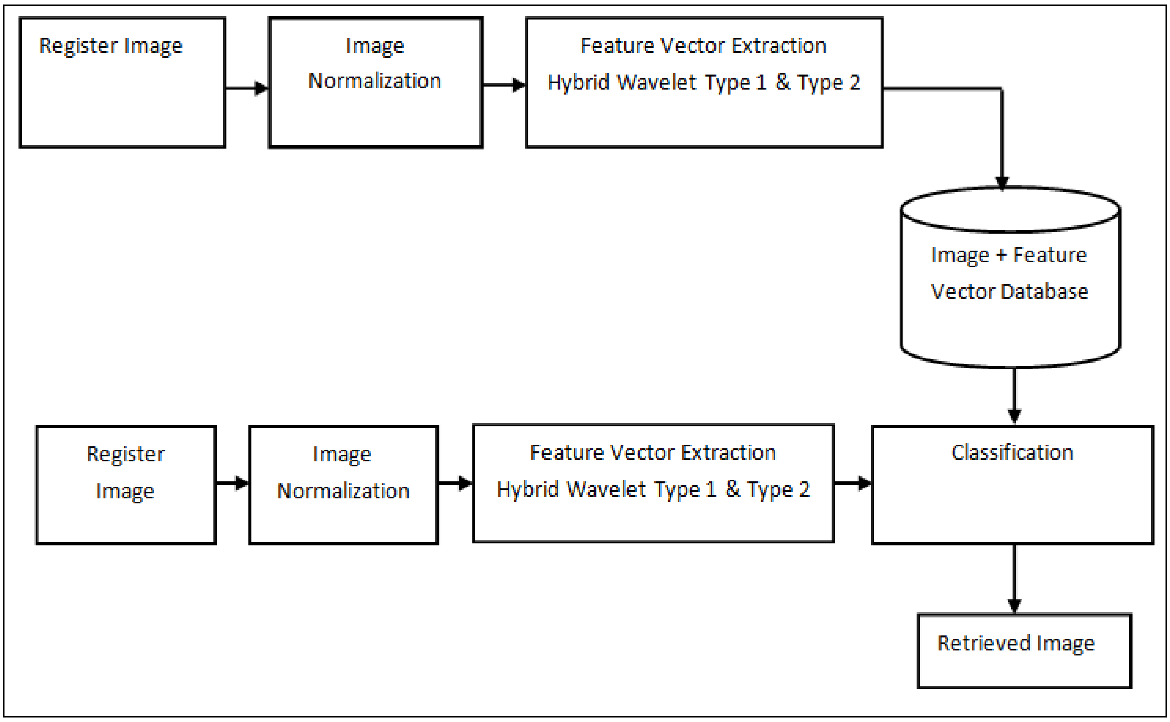


**Fig 2: Architecture of CBIR [10**

### 4.1 Uploading Operation

The Upload operation consists of uploading the image and the feature vector in the blob and table storage on the cloud using a web service running on the web role. Basic indexing of the image is done at the stage of Uploading. It also consists of background processing of feature extraction done on the image at the web role. The techniques used for extracting feature vectors are Hybrid wavelet type1 and type2.

The uploading operation consists of the following steps shown in figure (3):

1. User will upload image on web service running on Web role.
2. Web role will compute feature vector of image using Hybrid Wavelet Type 1 and Type 2.
3. Web role will upload image as well as feature vector into blob storage.
4. Once image is being uploaded in blob storage an entry will be made in table storage with its ID and location of generated feature vector.

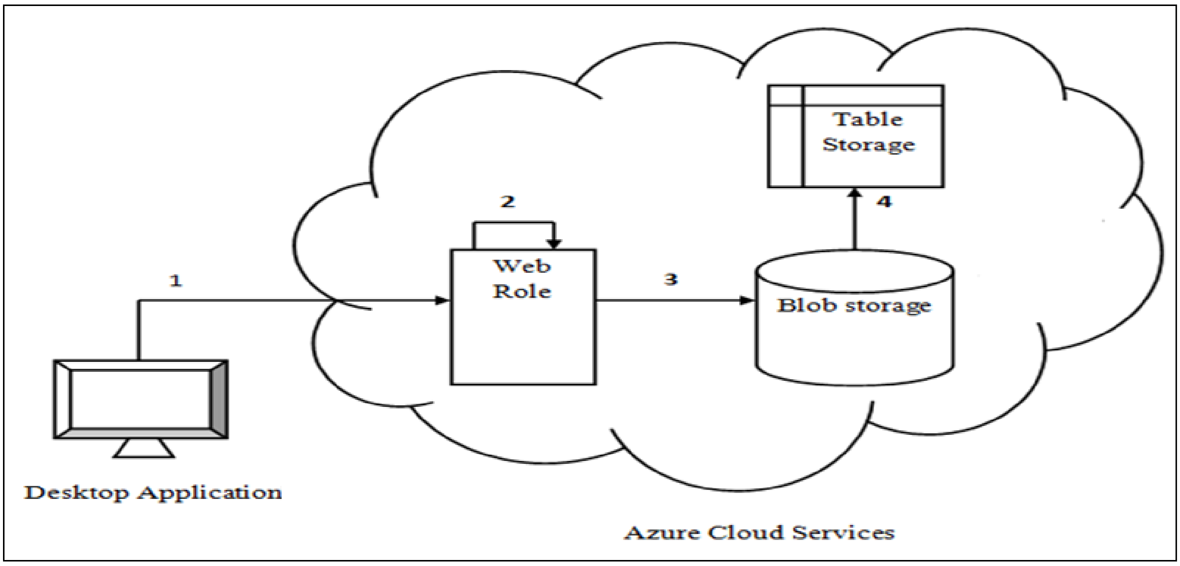


**Fig 3: Uploading Operation**

## 4.2 Searching Operation

The Search operation consists of verification and retrieval of the images in the database. This is done by computing/extracting the feature vector of query image and comparing this computing/extracted feature vector with the feature vectors stored in the blob storage. If the calculated feature vector matches with the feature vector in the blob storage with the high value than the threshold, then image will be delivered from the database.

The Searching Operation consists of the following steps shown in figure (4):

1. The image to be searched will be first uploaded on web service running on web role.
2. Web role will compute feature vector of image using Hybrid Wavelet Type 1 and Type 2.
3. Web role will upload image to be searched as well as feature vector into blob storage.
4. Once image is being uploaded in blob storage web role will get the location of the feature vector from the blob storage.
5. Location of the matched feature vector will be given to the web role.
6. Web role will go to the location in Blob storage.
7. Web role will get the matched image from that location.
8. The web role will give back the retrieved images onto the user screen.

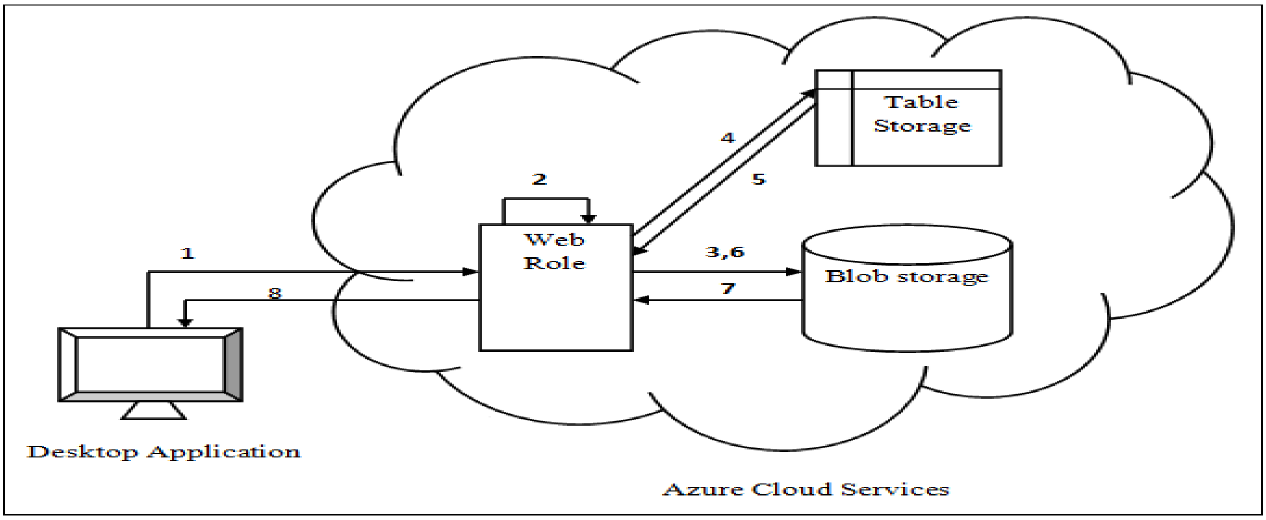


**Fig 4: Searching Operation**

## 5. RESULTS

The results include the images browsed from the local directory and uploading this image in the blob storage on the cloud.

## 5.1 Image Browsed from Local Directory

The image is browsed by a visual basic template. The Figure 5 shows the visual basic template to browse image from the local directory. A class is selected, so that the browsed image will be stored in that respective image class.

## 5.2 Uploading image to the cloud

For uploading the image to the cloud, a storage should be created on the cloud first. In this case, windows azure storage services are used to create a cloud storage. Once the windows azure has created the storage it provides an indication on its management portal. Now once we click on the select image to upload button given in the image browsing form, image will be moved to cloud storage as shown in Figure 6 and Figure 7.

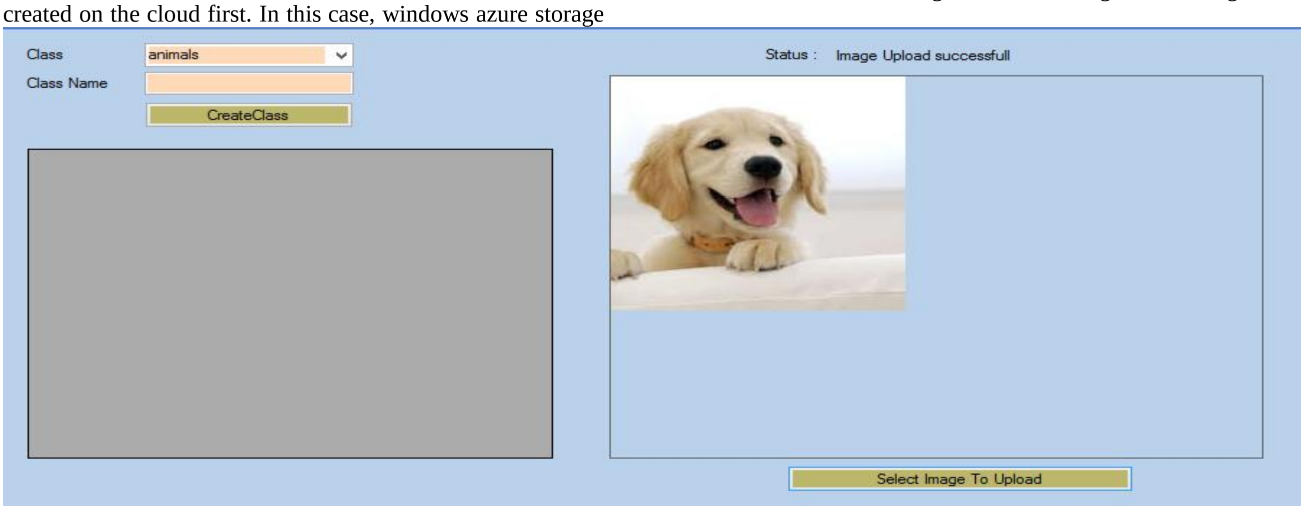


**Fig 5: Image browse Form**

| NAME | URL | LAST MODIFIED |
|---|---|---|
| animals | https://cbirtestdb.blob.core.windows.net/animals | 2/22/2015 10:57:34 PM |
| antiques | https://cbirtestdb.blob.core.windows.net/antiques | 2/23/2015 10:01:43 AM |
| bikes | https://cbirtestdb.blob.core.windows.net/bikes | 2/22/2015 8:40:48 PM |
| birds | https://cbirtestdb.blob.core.windows.net/birds | 2/22/2015 8:50:11 PM |
| clothes | https://cbirtestdb.blob.core.windows.net/clothes | 2/22/2015 8:51:56 PM |
| flowers | https://cbirtestdb.blob.core.windows.net/flowers | 2/22/2015 11:06:57 PM |
| food | https://cbirtestdb.blob.core.windows.net/food | 2/22/2015 8:54:01 PM |
| planes | https://cbirtestdb.blob.core.windows.net/planes | 2/22/2015 8:40:09 PM |

**Fig 6: Image Containers**

| NAME | URL | LAST MODIFIED | SIZE |
|---|---|---|---|
| E:/LocalFolder/10.jpg | https://cbirtestdb.blob.core.windows.net/animals/E:/LocalFolder/ | 2/23/2015 10:39:32 AM | 3.3 KB |
| E:/LocalFolder/11.jpg | https://cbirtestdb.blob.core.windows.net/animals/E:/LocalFolder/ | 2/23/2015 10:39:44 AM | 5.81 KB |
| E:/LocalFolder/12.jpg | https://cbirtestdb.blob.core.windows.net/animals/E:/LocalFolder/ | 2/23/2015 10:39:50 AM | 3.55 KB |
| E:/LocalFolder/3.jpg | https://cbirtestdb.blob.core.windows.net/animals/E:/LocalFolder/ | 2/22/2015 10:57:41 PM | 192.99 KB |
| E:/LocalFolder/4.jpg | https://cbirtestdb.blob.core.windows.net/animals/E:/LocalFolder/ | 2/23/2015 10:38:52 AM | 32.35 KB |
| E:/LocalFolder/5.jpg | https://cbirtestdb.blob.core.windows.net/animals/E:/LocalFolder/ | 2/23/2015 10:39:02 AM | 6.59 KB |
| E:/LocalFolder/6.jpg | https://cbirtestdb.blob.core.windows.net/animals/E:/LocalFolder/ | 2/23/2015 10:39:08 AM | 4.79 KB |
| E:/LocalFolder/7.jpg | https://cbirtestdb.blob.core.windows.net/animals/E:/LocalFolder/ | 2/23/2015 10:39:12 AM | 4.38 KB |
| E:/LocalFolder/8.jpg | https://cbirtestdb.blob.core.windows.net/animals/E:/LocalFolder/ | 2/23/2015 10:39:17 AM | 4.52 KB |
| E:/LocalFolder/9.jpg | https://cbirtestdb.blob.core.windows.net/animals/E:/LocalFolder/ | 2/23/2015 10:39:23 AM | 4.9 KB |

**Fig 7: Uploaded Image Files**

## 6. CONCLUSION

Thus this paper describes how the images will be browsed from local directory, and how they will be stored in a blob storage on cloud. CBIR SaaS architecture is proposed due to which the services of CBIR will be dynamically made available throughout the desired systems resulting in increase in applications scalability, flexibility and availability.

Performance will be evaluated using Similarity measures like precision and recall. The outcome of the similarity measure is expected to be higher than the threshold value. The cloud services provided by cloud architecture will handle all the unexpected traffic, and it will simultaneously benefit with minimized cost. CBIR will no longer behave as a product and hence will be available to the intended users dynamically.

## 7. REFERENCES

[1] A. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain (2000), "Content-based image retrieval at the end of the early years", in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, 1349–1380, December 2000.

[2] Yogita Mistry and D.T. Ingole, "Survey on Content Based Image Retrieval Systems" in *International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)*, Vol.1, October 2013.

[3] Zhuo YANG, Sei-ichiro KAMATA and Alireza AHRARY, "NIR: Content Based Image Retrieval on Cloud Computing" in *IEEE International Conference on Intelligent Computing and Intelligent Systems*, 2009.

[4] Zainab Ibrahim Abood, Israa Jameel Muhsin, Nabeel Jameel Tawfiq, "Content-based Image Retrieval (CBIR) using Hybrid Technique" *in International Journal of Computer Applications (0975 – 8887)*, December 2013.

[5] H B Kekre, V A Bharadi, S D Thepade, B K Mishra, S E Ghosalkar, S M Sawant, "Content Based Image Retrieval Using Fusion of Gabor Magnitude and Modified Block Truncation Coding" in *IEEE computer society*, 2010.

[6] Ja-Hwung Su, Wei-Jyun Huang, Philip S. Yu and Vincent S. Tseng, "Efficient Relevance Feedback for Content-Based Image Retrieval by Mining User Navigation Patterns", in *IEEE Transactions On Knowledge And Data Engineering*, Vol. 23, No. 3, 1041-4347/11, March 2011.

[7] Arti Khaparde, B L Deekshatulu, M.Madhavilatha, Zakira Farheen, Sandhya Kumari, "Content Based Image Retrieval Using Independent Component Analysis" in *IJCSNS International Journal of Computer Science and Network Security*, VOL.8 No.4, April 2008.

[8] S. Selvarajah and S. R. Kodithuwakku "Combined Feature Descriptor for Content Based Image Retrieval", 2011 6th International Conferencon Industrial and Information Systems, ICIIS 2011, Aug. 16-19, 2011, Sri Lanka, IEEE.

[9] M. Arevalillo-Herr´aez, Francesc J. Ferri, Salvador Moreno-Picot, "An interactive evolutionary approach for content based image retrieval", in *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, San Antonio, TX, USA - October 2009.

[10] A R Sawant, V A Bharadi, H B Kekre, Bijith Markarkandy, "Implementation of Content based Image Retrieval and Comparison using Different Distance Measures" *in IJAIS – ISSN : 2249-0868 Foundation of Computer Science FCS, New York, USA*, (ICWAC) 2013.

[11] Cloud Service: http://azure.microsoft.com/en-us/documentation/articles/cloud-services-what-is/

[12] Azure Storage: http://azure.microsoft.com/en-in/documentation/articles/storage-introduction/

[13] Blob Storage: http://azure.microsoft.com/en-us/documentation/articles/storage-dotnet-how-to-use-blobs/