International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 08 | August -2020 ISSN: 2582-3930

Content based Image retrieval using FAST Machine learning approach in Cloud computing

Abhishek S G
Computer Science dept.
Cambridge Institute of Technology
Bengaluru, India

Dishant Kedia
Computer Science dept.
Cambridge Institute of Technology
Bengaluru, India

Hardika K
Computer Science dept.
Cambridge Institute of Technology
Bengaluru, India

Hema M
Computer Science dept.
Cambridge Institute of Technology
Bengaluru, India

Abstract –**The images are having significant role in our daily life. Images consume more storage space when compared to text documents. For preserving privacy of images, before deploying it to cloud storage images are encrypted. A scheme supporting CBIR (content-based image retrieval) from encrypted images is proposed in this paper. The features are identified from the outsourced images and by applying locality-sensitive hashing pre-filter tables are generated for increasing the efficiency of searching. The features of the outsourced images are represented by using interest points and are encrypted by using a stream cipher. A machine learning algorithm using FAST method identifies the interest point on image contour, which helps in retrieving most similar images from cloud. Besides these, for avoiding illegal distribution of retrieved images by query, the cloud server embedds a unique watermark to the encrypted images by using a watermark based protocol. The average search time and precision of the proposed system can be inferred from performance evaluations.**

Keywords– *Cloud computing, Content-based image retrieval (CBIR), Contour-based shape descriptor, Machine learning, Watermark Features from accelerated segment test (FAST)*

I. INTRODUCTION

Image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Most traditional and common methods of image retrieval utilize some method of adding metadata such as captioning, keywords, or descriptions to the images so that retrieval can be performed over the annotation words. Manual image annotation is time-consuming, laborious and expensive; to address this, there has been a large amount of research done on automatic image annotation. Additionally, the increase in social web applications and the semantic web have inspired the development of several web-based image annotation tools. In Boosting Image retrieval, the approach is predicated on the assumption that each image is generated by a sparse set of visual “causes” and that images which are visually similar share causes. There is a mechanism for computing a very large number of highly selective features, which capture some aspects of causal structure. At query time a user selects a few example images, and a technique known as “boosting“ is used to learn a classification function in this feature space. In soft query in image retrieval system, the use of soft computing and user defined classifications in multimedia database systems for content-based queries is explored.

With traditional database systems, objects/tuples are grouped into classes/relations using “hard" membership. Hence, the result of a query to obtain the members of a class is a fixed set. With multimedia databases, however, an object may belong to different classes with different probabilities (“soft" membership). The model is implemented by extending the traditional database query capabilities such that the result of a query depends on the user who submits the query. Content-based image retrieval (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision techniques to the image retrieval problem, that is, the problem of searching for digital images in large databases. "Content-based" means that the search will analyze the actual contents of the image rather than the metadata such as keywords, tags, and/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. CBIR is desirable because most web based image search engines rely purely on metadata and this produces a lot of garbage in the results. The general image retrieval system (Fig. 1) consists of three main modules such as input module, query module, and retrieval module [2]. In the input module, the feature vector is extracted from input image. It is then stored along with its input image in the image database. On the other hand, when a query in the input module, the feature vector is extracted from input image. It is then stored along with its input image in the image database. On the other hand, when a query image enters the query module, it extracts the feature vector of the query image. In the retrieval module, the extracted feature vector is compared to the feature vectors stored in the image database. As a result of query, the similar images are retrieved according to their closest matching scores. Finally, the target image will be obtained from the retrieved images.

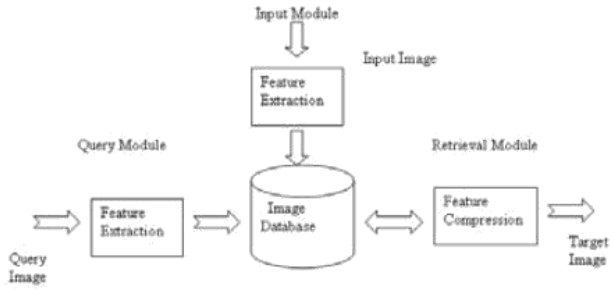


Fig 1. Retrieval of target image

II. PROPOSED SYSTEM

We are developing a image retrieval and question answering system based on captioning. Using the Content-based image retrieval (CBIR) technology, in which we ensure the searching for digital images in large databases in encrypted format and retrieving them. Our project includes following features:

✓ Images are encrypted before being uploaded into the cloud. AES algorithm is used in this project for encryption and decryption purpose.

✓ The Haar transform is used for image compression. The basic idea is to transfer the image into a matrix in which each element of the matrix represents a pixel in the image. We can keep track of all the sensor reading through android app.

✓ Watermark Based Protocol is applied on the decrypted image to avoid illegal distribution

✓ Features from Accelerated Segment Test (FAST). FAST is an interest point identification method using machine learning approach.

✓ locality-sensitive hashing is an algorithmic technique that hashes similar input images into the same "buckets" with high probability. Since similar images end up in the same buckets, this technique is used for data clustering.

✓ Machine Learning algorithm-KMeans is implemented in which the Calculated index is put into clusters so that the searching of images is efficient.

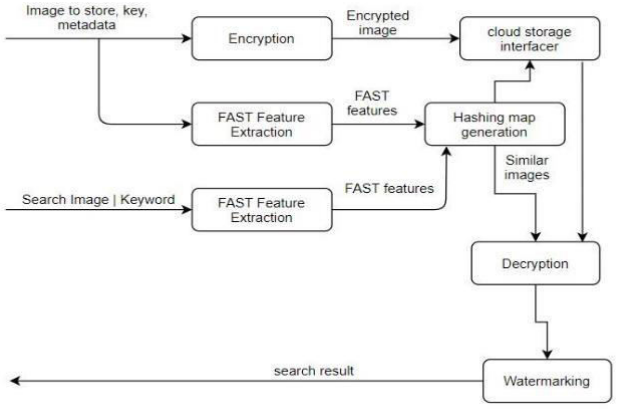


Fig 2. System working

Owner who want to store image for retrieval provides the image, key and meta data information. FAST features are extracted from the image, the image and the meta data is then encrypted and stored into cloud. The hash key is generated from the FAST features and added to the hash map. When user wants to lookup for the images, he provides the image or keyword. If the image is given as input FAST features are extracted from the image , hash key is generated from the FAST features and lookup is done of the hash map to get the similar images. If keyword is given, it is encrypted and looked for match in the encrypted meta information in the cloud for the matching images. If the user provides the key, the matching images are decrypted and watermarked. The watermarked image is returned as the result.

III. SYSTEM ARCHITECTURE

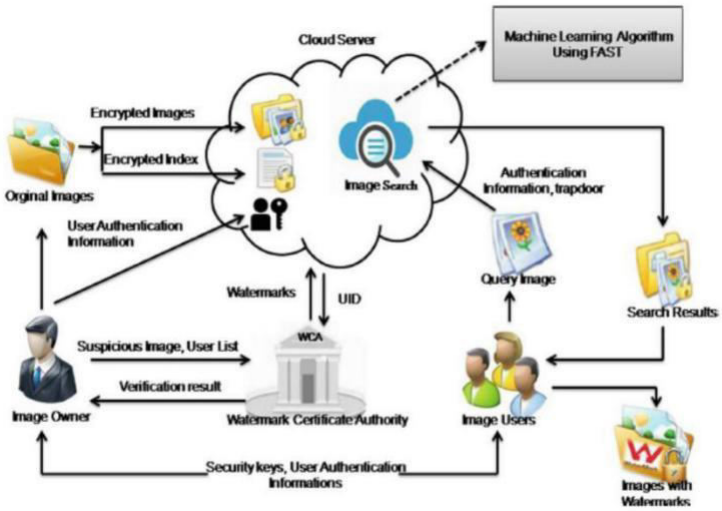


Fig 3. Framework of the CBIR from cloud computing using FAST Machine Learning Algorithm

Image Owner is authorized user who can deploy his images to the cloud server in an encrypted form. The values of features of an image, F are obtained from images M, and I, a secure index is generated using F. The collection of encrypted images C, encrypted index I and the identity of the image owner are sent to the server for storage.

Image users can retrieve required images from cloud server by using query image. The retrieval of images is achieved by using a trapdoor (TD). After receiving requested image, a secret key is shared by the image owner to the user for decryption.

Cloud server consist of the encrypted images (C) and the encrypted index (I) for each image outsourced by image owners. The query requests from users are solved and besides these the cloud server takes the duty of watermarking the requested images.

WCA is a trusted authority who helps in generating watermarks for the image users by using watermarking algorithm

A class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a

system by showing the system's classes, their attributes, and the relationships between the classes.

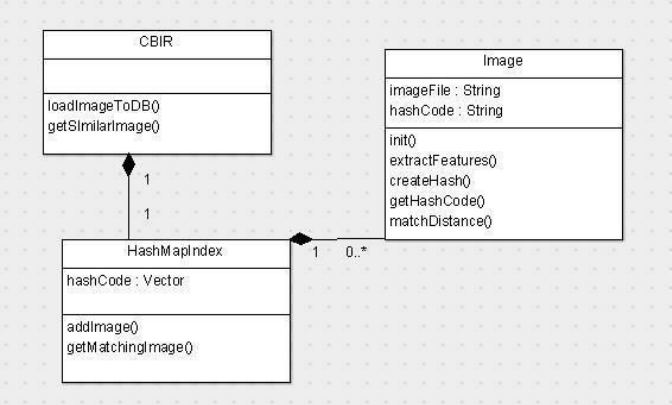


Fig 4. Class diagram

A use case diagram is a type of behavioral diagram created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases.

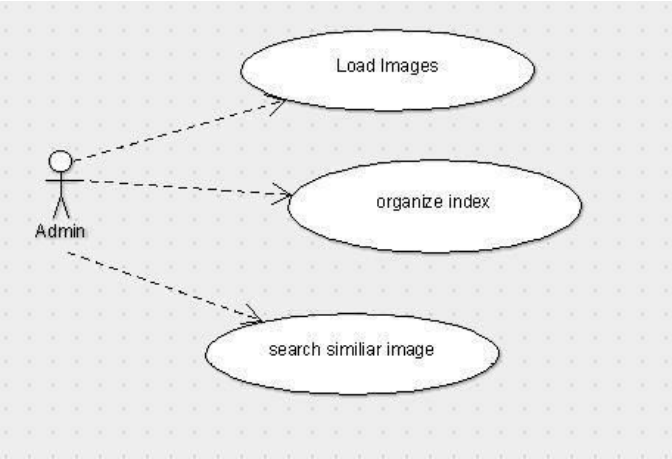


Fig 5. Use case

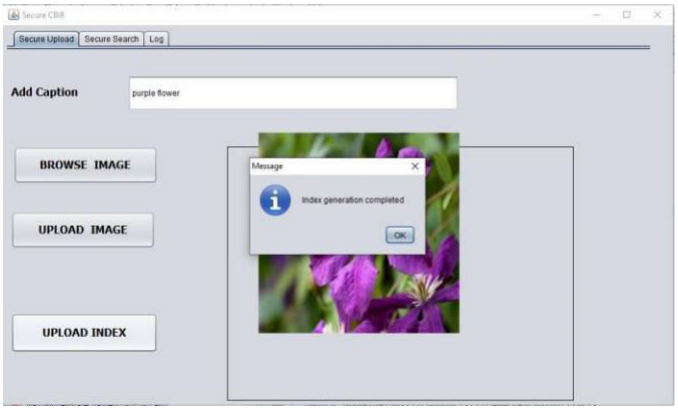
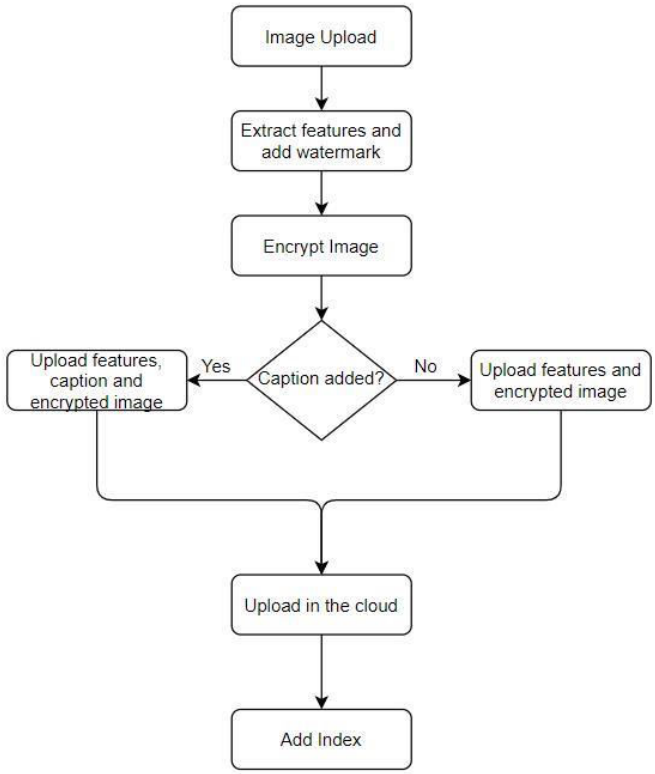


Fig 6. Index Upload

IV. METHODOLOGY

The proposed system is to design highly efficient search using images. It performs both updating and retrieving images from the cloud.

To update the cloud with an image, the image is uploaded from the local storage and then it is processed. Different features are extracted from the images using FAST and HAAR method. Watermark is applied to the images to avoid illegal distribution of images. Then the images are encrypted using AES algorithm. Captions can also be added to the images. The encrypted images along with the features and the caption (if specified) are uploaded to the cloud. After the image is uploaded in the cloud, index is assigned to it which provides the link between the image and its features.

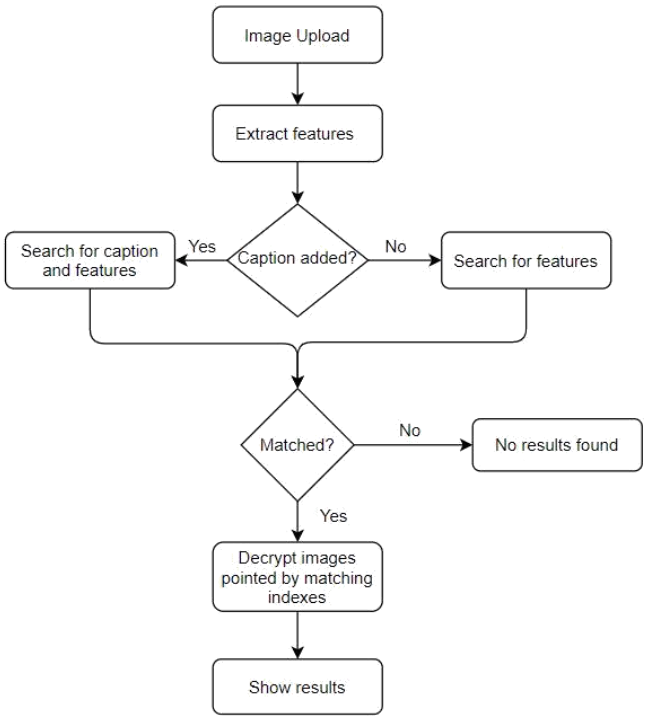


There are two ways to perform the search :

Caption Based Search
Image Based Search

In caption based search, the query caption is searched for in the cloud. Images pointed by the indexes of the similar captions are retrieved. The retrieved images are decrypted using the AES algorithm and are sent to the user.

In image based search, the query image is processed using a FAST and HAAR algorithm and the features are extracted. Search is performed in the cloud with the retrieved features. Images pointed by the indexes of the similar features are retrieved. The retrieved images are decrypted using the AES algorithm and are sent to the user. For more efficient results, both images and the captions are passed as the query.



V. RESULT

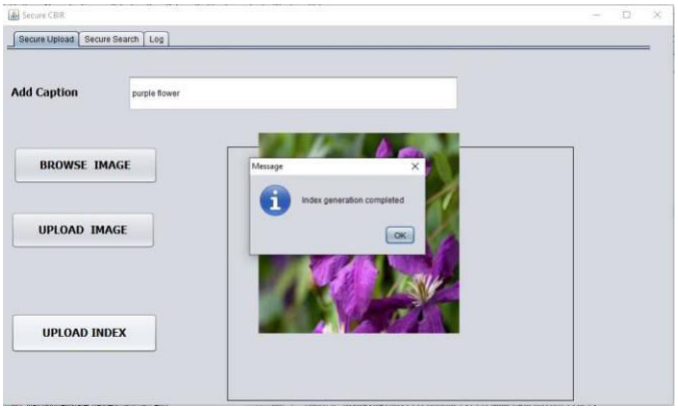


Fig. 7. Uploading image to cloud

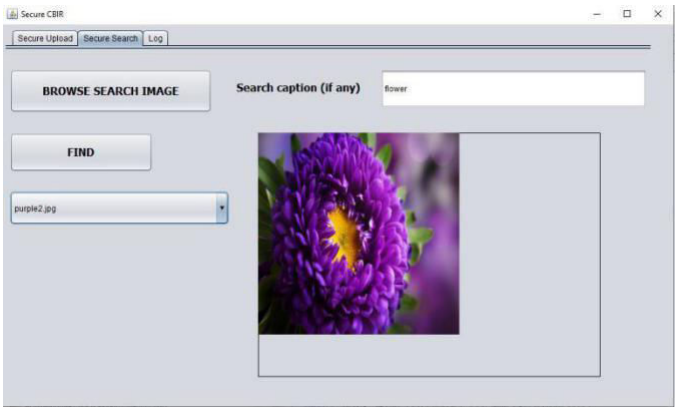


Fig. 8. Similar image retrieved

VI. CONCLUSION

This work provides a Image retrieval system using CBIR. The image features are represented by using interest points identified by FAST. The locality sensitive hashing is utilized to group images having similar feature values which improve the search efficiency. Then, the machine learning algorithm based on FAST is applied over the outsourced images for identifying the similar images. Based on these identified interest point values the similarity score is obtained and the cloud server rank images without much effort. Since FAST algorithm is implemented by identifying the interest points on the detected contour, query user can easily retrieve the most similar images having common features with a better search efficiency Integration of text based image retrieval along with content based image retrieval is done for better results. Security is provided for the search results using watermarking.

ACKNOWLEDGEMENT

We are so much thankful for our project guide, Dr Josephine Prem Kumar. who supported us during the project completion and also helped us to improve the manuscripts significantly.

REFERENCES

[1] N. Sharmi, P. Mohamed Shameem, and R. Parvathy "Content-Based Image Retrieval Using FAST Machine Learning Approach in Cloud Computing". Department of CSE, TKM Institute of Technology, Kollam,Kerala, (2018)

[2] Xia, Z., Wang, X., Zhang, L., Qin, Z.: A privacy-preserving and copy-deterrence contentbased image retrieval scheme in cloud computing. IEEE Trans. Inf. Forensics Secur. 11(11), 2594–2608 (2016),

[3] Lu, W., Swaminathan, A., Varna, A.L., Wu, M.: Enabling search over encrypted multimedia databases. In: Proceedings of SPIE, vol. 7254, p. 725418, February.

[4] Nishant Singh, Shiv Ram Dubey, Pushkar Dixit, Jay Prakash Gupta, "Semantic Image Retrieval by Combining Colour, Texture and Shape Features", International Conference on Computing Sciences,2012."

[5] Ela Yildizer, Ali Metin Balci, Mohammad Hassan, Reda Alhajj, "Efficient content-based image retrieval using Multiple Support Vector Machines Ensemble", Expert Systems with Applications 39, 2385–2396, 2012.

[6] Lijun Zhao, Jiakui Tang, "Content-Based Image Retrieval Using Optimal Feature Combination and Relevance Feedback", International Conference on Computer Application and System Modeling,volume-4,

[7] Bingxin Xu, Qian Yin, Guangjun Lv, "Using SVM to Organize the Image Database", International Conference on Computational Intelligence and Security, 2009.

[8] P. S. Hiremath , Jagadeesh Pujari, "Content Based Image Retrieval using Colour, Texture and Shape features", 15th International Conference on Advanced Computing and Communications,2007.