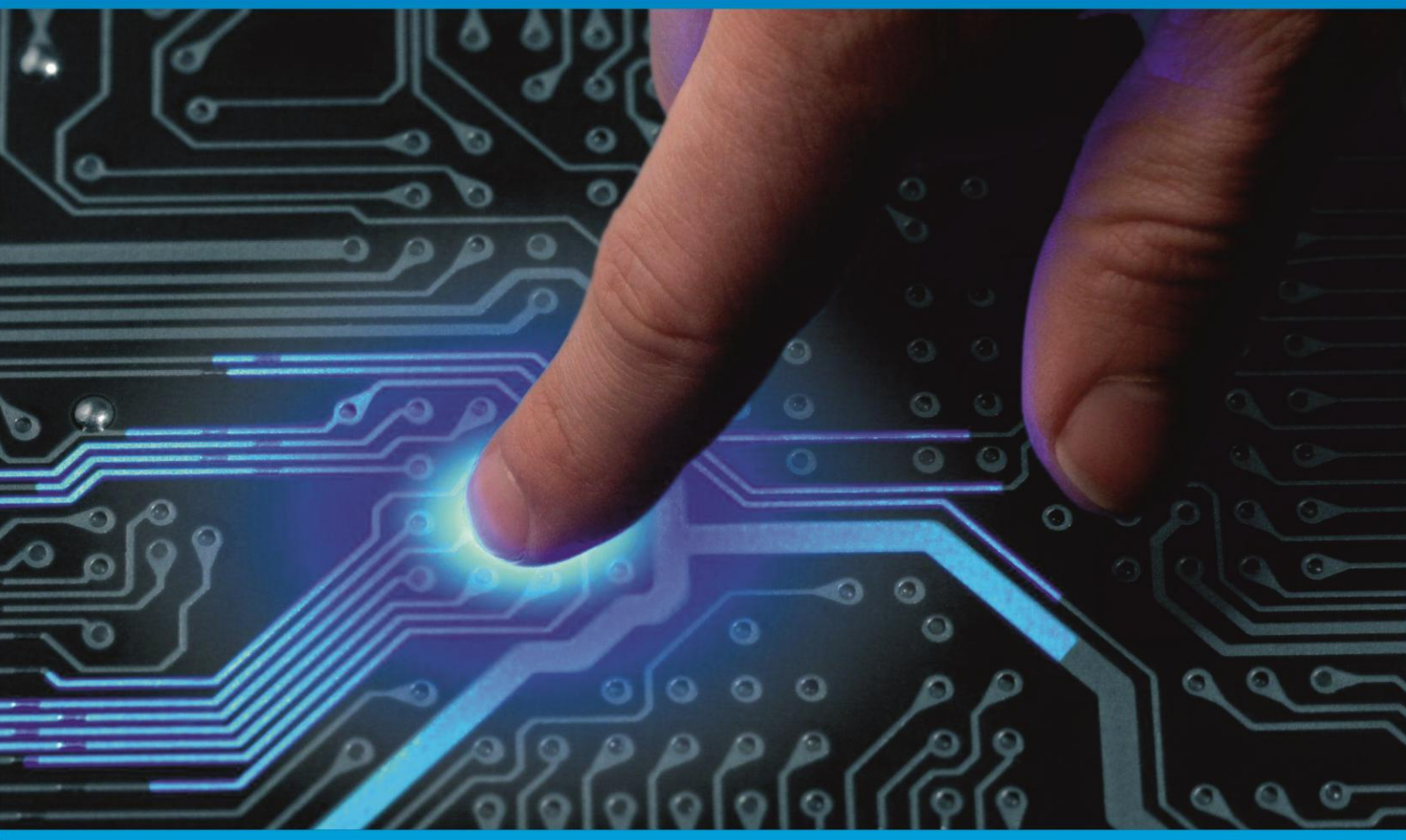
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 8, Issue 10, October 2020

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

|| Volume 8, Issue 10, October 2020 ||

An Enhancement of CBIR using LTrP with Hadoop

Prof. Vijaysinh G. Chavan[1], Pooja R. Gurav[2]

Assistant Professor, Dept. of CSE, SKN Sinhgad College of Engineering, Pandharpur, Maharashtra, India[1]

PG Student, Dept. of CSE, SKN Sinhgad College of Engineering, Pandharpur, Maharashtra, India[2]

ABSTRACT: Today, huge quantity of data, in the form of images, is produced through digital cameras, mobile phones and photo editing software. A large part of this data is stored in online repositories. This data is private to each user and consequently it should not be accessible by others. In earlier systems, images were searched by the tags or keywords or description assigned to them wherein if an image is wrongly described, querying image search will not result in the required image. Here, a 'content based' search has been proposed which analyzes the content of the image. System will allow user to upload a particular image and depending on combined values of color, shape and texture, the system will retrieve similar images from database. The system incorporates techniques for upload and search of images over large datasets of images based on the content of the images rather than the keywords and tags or any other textual information. Another part of the system is Hadoop distributed file system (HDFS). Hadoop defines a framework which allows processing on distributed large sets across clusters of computer.

KEYWORDS: Content Based Image Retrieval; LTrP; hadoop; cluster; texture;

I. INTRODUCTION

Nowadays, technology has become progressively advanced. This has led to cheap and advanced multimedia devices which have given rise to huge data volumes. This massive data need to be stored in the database for various applications such as prevention of crime, medical, security, etc. These applications have created a requirement for effective and efficient methods of storage, search and retrieval of images via parallel processing techniques. With growing technology, security has become a crucial concern for visual forms of information. For example, suppose an organization has developed a novel face recognition algorithm. The organization will wish that the input images as well as images in their database are not revealed publicly by the algorithm or any other means. We attempt to find a method for storage and retrieval of images in a way that such an objective is achieved. Earlier, the images were stored with associated labels, and searching was done on the basis of these labels. But such a method is prone to many errors. If the images are wrongly annotated, irrelevant images might be retrieved. Moreover, it is a laborious task to search for an image which has been assigned a wrong label [3].

Nowadays, with extensive reach of social networks, users have become more concerned about their privacy and their data being stored on servers. Some users even prefer to hide their details from database admin. If stored information on the server is seen by database admin, a user's privacy is breached and the chances of misuse of this information increase. If an organization's employee details are stored on a server along with their photographs for face recognition, the organization prefers to keep this information out of reach of any other user or database admin [1]. If this information is not hidden perfectly, a compromised database admin may be able to access them and use it for his/her own benefits. Such a situation needs to be avoided.

Content Based Image Retrieval

The term "Content-Based Image Retrieval" is used for retrieving the corresponding images from the database. Based on their feature of images which derived the image itself like texture, color and shape and domain specific like human faces and fingerprints. The retrieval on the based on the content of an image is to be more effective than the text based which is called content based image retrieval that are used for a various application like vision techniques of the computer [2]. Traditionally, search of the images are using text, tags or keywords or annotation assigned to the image while storing into the databases. Whereas if the image stored in the database is not uniquely or specifically tagged or wrongly described then it's insufficient, laborious and extremely time consuming job for search the particular image in the large set of databases [9]. For this purpose obtaining the most accurate result CBIR system are used which searches and retrieve the query images from the large databases based on their image content like color, texture and shape which derived from the image itself.

II. RELATED WORK

The private content of users, in form of images, should not be visible to any person except the owner. This data should even be protected from database admin. Images retrieval is increasing and crucial importance is given to domain specific information. Large and distributed collections of scientific, technical images are retrieved using sophisticated and precise measures of similarity and query based semantics [2]. The map reduces the framework to distribute the work in parallel and gives faster result in shorter time with the working on the Gigabyte and Petabyte of data. Image retrieval has been done on the combined value RGB values on the low, medium and large values whereas the images stored on the database are retrieved on the basis of neural methods [3]. S. Murala et al. proposed an algorithm for indexing and retrieval of images for Content Based Image Retrieval Using Local tetra Patterns (LTrPs). It defines the standard combination of local binary patterns (LBP) and local ternary patterns (LTP) which are computed by referenced pixel and surrounding its neighbors by calculating difference in gray level. The proposed method gives the relationship between referenced pixel with its neighbors pixel with respect to the direction that are computed by derivatives in first order in vertical and horizontal direction [4]. A. Hussain et al. stated that the importance of similarity measurement. The images need to be kept in databases due to increased quantity of digital data in fields like medicine, private life photos and journalism. In order to retrieve the desired images from the database, efficient and accurate retrieval system is required. An image can be retrieved based on the features like color, texture, shape and the content. Most similar images having the least distance are searched and given as output. [5].

III. PROPOSED METHODOLOGY

The developed system provides the application interface for users. The system incorporates two phases – 1) Upload the images and 2) Search the image from the database. Both modules are deployed on Hadoop cluster. For uploading phase, thedata is inputted for storage of HDFS and Query Image is searched from the database.

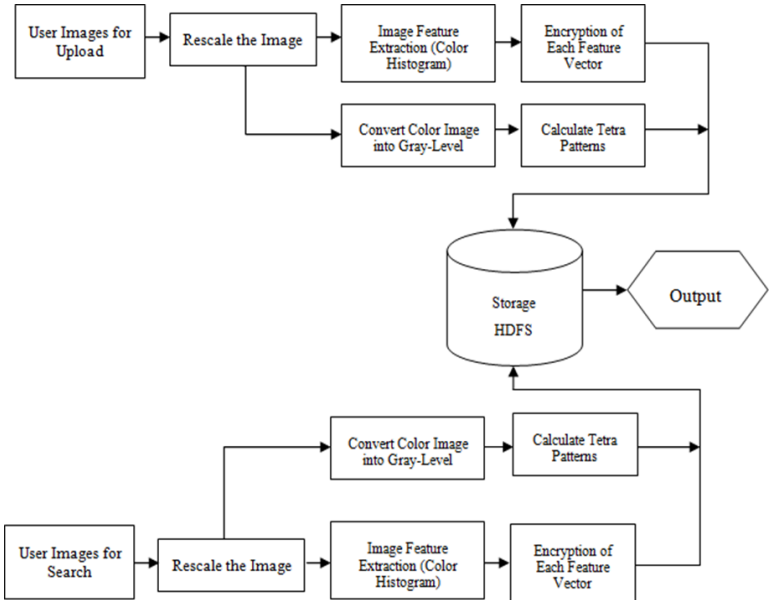


Figure 1: Flow Chart for Enhanced CBIR Using LTrP on Hadoop

A. *Description of the Proposed System:*

Upload Images:

System allows user to upload one or more images at a single point into the system. In this proposed system, adding the image into the databases has sub processes. The user uploads the images on database of Hadoop which is called HDFS. This process has been split into two phases as shown in figure 4.1.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

|| Volume 8, Issue 10, October 2020 ||

Phase 1:

User can upload one or more images into the database by giving the image path at a single point from the GUI. The uploaded images by the user may be of different sizes. Large or different sized images may take more time for computation. To deal this scenario, the system has the facility of rescaling the image. After the image has been rescaled, it is converted into grayscale. The first-order derivatives of vertical and horizontal axes are applied and every direction of the pixel that is calculated based on the center pixel forms patterns which are divided into four parts. Then tetra patterns are calculated and are separated into three binary patterns. Magnitude of the center pixel is computed. The feature extracted from the binary patterns and binary patterns obtained from the magnitude patterns are combined to form a feature vector. The converted image is then stored in .text file format on the database.

LTrP:

Local Tetra Patterns are the combination of LBP, LDP and LTP. It describes the special local texture structure using center gray pixel by direction. The center level direction of the gray level pixel is denoted by I. Denotes the center pixel in I, h horizontal and vertical neighborhoods of respectively. Each part is converted into three binary patterns by tetra patterns. [1]

$$I0°1(gc)=I(gh)-I(gc)$$
$$I90°1(gc)=I(gv)-I(gc)$$
$$MI1(gp)=\sqrt{I0°1(gp))2+I90°1(gp))2}$$

Phase 2:

After the image is rescaled, as described in phase 1, its color features are calculated i.e. RGB values for each pixel of image. These calculated feature values are encrypted with simple encryption and a .txt file is formed from the result which is then stored into the database.

Searching of Images:

Like uploading phase, system provides GUI for user to search and retrieve the images from the database by the query image. When a user wishes to search a particular image, he/she has to provide the search image path. The processes included in upload phase are same for search process. The searched image is matched with the help of pixel value for LTrP method. It is also searched by color method where the color encrypted values are compared with the values stored in the database.

IV. RESULTS

The effect of enhancement in CBIR by using Local Tetra Patterns using Hadoop desktop computer system. The system outcome shows improvement in the basic CBIR.Initially the experiments are carried on the small set of images for evaluation and testing. The code was run on this set and all operations were performed on stored .txt files. This system allows images of a variety of sizes and formats. For testing and results, a face dataset has been used. The images, in the dataset, are in .jpg format and have the dimension of 640 by 480 pixels. Experiments are carried out at upload and search stage with repository of images.

Experiments are carried out on Hadoop cluster which has one master PC and two slave PCs. Each of the PC has different processor, RAM and disk storage. The data to be stored on the HDFS are inputted with the help of GUI. We have tested uploading time for images from one image to 1500 images at one point and search a particular image on database. The time taken for upload and search are measured in milliseconds.

System Implemented	Security	Upload Process	Color Histogram	LTrP Method	Metadata Search
Ostensive Model(OM)	No	No	Yes	No	No
FIRE	No	No	Yes	No	No
Triangle Inequality Algorithm(TIA)	No	No	-	-	No
Imgseek	No	No	Yes	No	No
My system	Yes	Yes	Yes	Yes	No

V. CONCLUSION AND FUTURE WORK

We experimented and evaluated the proposed system with four aspects. One is uploading, second is search query image by CBIR Method, third is the query image search by the color values and fourth is to store all encrypted image value to database which provides security.Images are growing through the various digital devices and these images are added to the image databases and internet for various applications. These images need to be stored and retrieved in effective and efficient manner. The searching time is the most important for any search method while searching it in large datasets of images. For these purposes, we described the novel method CBIR and color histogram for PCBIR of large datasets incorporated with HDFS file system of Hadoop framework.

This application is designed by combining a number of different domains into one. After implementing this system, we find that there is wide scope for different feature extractions like shape and for secure storage of images on database using different encryption techniques. This application has large scope in cloud domain and internet world because every domain servers are using cloud technology and they want to provide their customers with new cutting edge applications for image storage and retrieval which are secure, efficient and are delivered with high throughput.

REFERENCES

[1] J. Shashank, "Content Based Image Retrieval Using Color, Texture and Shape Features," International Conference on Advanced Computing and Communications, ADCOM , pp.780 – 784,2007.
[2] Shankar M. Patil "Content Based Image Retrieval Using Color, Texture and Shape," International Journal of Computer Science & Engineering Technology (IJCSET), Vol. 3, Sept. 2012.
[3] Sitalakhmi and S. kulkarni "MapReduce neural network framework for efficient content based image retrieval from large datasets in the cloud" IEEE Transactions on Hybrid Intelligent Systems, pp. 63 – 68,2012.
[4] Ryszard S. Choras "Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems" International Journal of Biology and Biomedical Engineering, Vol. 1, 2007.
[5] Murala S., Maheshwari R.P., and Balasubramanian R., "Local Tetra Patterns: A New Feature Descriptor for Content-Based Image Retrieval," IEEE Transactions on Image Processing, Vol. 21,pp.2874 – 2886, May 2012.
[6] Khan, S.M.H. ,Hussain, A. , and Alshaikhli, I.F.T."Comparative Study on Content-Based Image Retrieval (CBIR)" International Conference on Advanced Computer Science Applications and Technologies (ACSAT), 2012 .
[7] Oberoi Ashish, BakshiVarun, Sharma Rohini and Singh Manpreet "A Framework for Medical Image Retrieval Using Local Tetra Pattern," International Journal of Engineering Science & Technology, Vol. 5, pp.27, Feb2013.
[8] Liangliang Shi , Bin Wu ,Bai Wang and Xuguang Yan "Map/reduce in CBIR application," International Conference on Computer Science and Network Technology (ICCSNT), Vol.4 , pp. 2465 – 2468, Dec. 2011.Chapter 7.Textures [Online].
[9] Muneto Yamamoto and Kunihiko Kaneko, "Parallel Image Database Processing With MapReduce And Performance Evaluation In Pseudo Distributed Mode,"
[10] International Journal of Electronic Commerce Studies,Vol.3, No.2, pp.211-228, 2012.
[11] Apache hadoop. [Online]. Available: http://hadoop.apache.org/
[12] Mapreduce - hadoop wiki. [Online]. Available: http://wiki.apache.org/ hadoop/MapReduce.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

|| Volume 8, Issue 10, October 2020 ||

[13] Hdfs users guide. [Online]. Available: http://hadoop.apache.org/docs/ hdfs/current/hdfs userguide.html.
[14] Apache hbase. [Online]. Available: http://hbase.apache.org/
[15] P. R. Sabbu, U. Ganugula, S. Kannan, and B. Bezawada, “An oblivious imageretrieval protocol," in Proceedings of the 2011 IEEE Workshops of International Conference on Advanced Information Networking and Applications, ser. WAINA'11. Washington, DC, USA: IEEE Computer Society, pp. 349{354}, 2011.
[16] Y. Rui and T. S. Huang, “Image retrieval: Current techniques, promising direc- tions and open issues," Journal of Visual Communication and Image Representation,vol. 10, pp. 39{62, 1999.
[17] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani,J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, “Query by image and video content: the QBIC system," Computer, vol. 28, no. 9, pp. 23{32, 1995.
[18] C.-C. Chen and H.-T. Chu, “Similarity measurement between images,"in Proceedings of the 29th annual international conference on Computer software and applications conference, ser. COMPSAC-W'05. Washington,DC, USA: IEEE Computer Society, pp. 41{42, 2005.
[19] C. W. Niblack, R. Barber, W. Equitz, M. D. Flickner, E. H. Glasman, D. Petkovic,P. Yanker, C. Faloutsos, and G. Taubin, “QBIC project: querying images by content, using color, texture, and shape," C. W. Niblack, Ed., vol. 1908, no. 1.SPIE, pp. 173{187, 1993.
[20] P. R. Sabbu, U. Ganugula, S. Kannan, and B. Bezawada, “An oblivious image retrieval protocol," in Proceedings of the 2011 IEEE Workshops of International Conference on Advanced Information Networking and Applications, ser. WAINA'11. Washington, DC, USA: IEEE Computer Society, 2011.

INNO SPACE

SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details