iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VIII **Month of publication:** August 2021

DOI: https://doi.org/10.22214/ijraset.2021.37852

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VIII Aug 2021- Available at www.ijraset.com

A survey on Content Based Image Retrieval in Cloud Environment with Privacy Preservation & Copy Deterrence

Monika Bansode[1], Dr. Deepa Deshpande[2]
[1]*Student of MGM's Jawaharlal Nehru Engineering College, Aurangabad*
[2]*Professor (CSE Department), MGM's Jawaharlal Nehru Engineering College, Aurangabad*

Abstract: ***Importance of images in day to day life increased tremendously. Therefore Content Based Image Retrieval studied extensively. Cloud computing offers on demand services to cloud user therefore many organizations prefer to use cloud for data storage. To protect images with sensitive or private information needs to be encrypted before being outsourced to cloud. However, this causes difficulties in image retrieval and management. The purpose of this study is to provide privacy preservation and copy deterrence Content Based Image Retrieval method using Lucene Indexing.***
Keywords: CBIR, Lucene Indexing, Copy Deterrence.

I. INTRODUCTION

Image data is part of one of the largest global Internet access situations in enterprise and personal use. Content based image retrieval, also commonly known as image search engine, one of the emerging technologies, attracts more and more people from different fields, such as computer vision, information retrieval, database systems, machine learning.

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.

Demand for efficient storage and retrieval services increases with the addition of a large database of images in all areas. While after more than twenty years of development, Content Based Image Retrieval techniques showcase the potential of many real-world applications. For example, doctors can use CBIR to find similar patients and facilitate clinical decision-making.

Large image databases often contain millions of images. CBIR services are often complex to store and calculate. Cloud computing is an excellent opportunity for access to resources, computing and extensive data storage, which is an attractive alternative for storing images and images employing CBIR. By outsourcing CBIR services to the cloud server, data owners are relieved of the need to maintain a local image database and interact with users.

Privacy is becoming a major concern of CBIR outsourcing. For example, patients may not want to disclose their medical images to others unless the CBIR's specialized physician uses it to determine the problem. Content-based image retrieval, 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 (see this survey[1] for a recent scientific overview of the CBIR field).

Content-based image retrieval is opposed to traditional concept-based approaches (see Concept-based image indexing). "Content-based" means that the search analyzes the contents of the image rather than the metadata such as keywords, tags, 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 searches that rely purely on metadata are dependent on annotation quality and completeness.

Having humans manually annotate images by entering keywords or metadata in a large database can be time consuming and may not capture the keywords desired to describe the image. The evaluation of the effectiveness of keyword image search is subjective and has not been well-defined.

In the same regard, CBIR systems have similar challenges in defining success. "Keywords also limit the scope of queries to the set of predetermined criteria." and, "having been set up" are less reliable than using the content itself.

II. LITERATURE REVIEW

A. *"A Privacy-preserving and Copy-deterrence Content-based Image Retrieval Scheme in Cloud Computing" IEEE Transaction (2016):*

In this two privacy threats are described as curious cloud serve look for additional information in owner's database and query user may illegally distribute the image to unauthorised user. Four visual descriptors are used and for feature vector encryption KNN is used. This is the first research to introduce searchable encryption scheme for dishonest query user for copy deterrence purpose. Author designed watermark based protocol, after completion of search operation watermark embedded to retrieved image. Author used watermark protocol for copy deterrence purpose, but it is not efficient. It requires high computational complexity and expected to be completed by cloud server. Slandering query user by embedding watermark related to the user in original image may be practice by image owner, the protocol proposed in this paper needs to prevent this type of illegal behaviour. Before illegal distribution query user may change the watermark image by image processing operations, so watermark bits should not be extracted with 100% accuracy. In this scheme only single image owner is considered. Proposed scheme performance depends on several parameters including the parameters in hash function, number of connected LHS function, and number of pre-filtered table and dimensionality of visual descriptor.

This scheme supports CBIR over encrypted images without leaking the sensitive information to the cloud server. Firstly, feature vectors are extracted to represent the corresponding images. After that, the pre-filter tables are constructed by locality-sensitive hashing to increase search efficiency, the feature vectors are protected by secure kNN algorithm, and image pixels are encrypted by a standard stream cipher. The similarity scores can be directly calculated with the encrypted features by the cloud server, which enables the cloud server to rank the images without the additional communication burden. The locality sensitive hashing is utilized to improve the search efficiency. Overall, the image features are secure against Cipher text-only Attack model, the image contents are secure against Chosen-plaintext Attack model, and the search efficiency is improved from $O(n)$ to $O(n')$. In this scheme there are some aspects that should be improve the proposed watermarking method is not very robust one and limited number of available watermarks.

B. *"Practical Privacy-Preserving Content-Based Retrieval in Cloud Image Repositories" IEEE 2016*

This paper is based on IES-CBIR, Image Encryption scheme with content based image retrieval properties. In IES-CBIR author made following security oriented measures, protection of image content is biggest concern that made by encrypting texture information with probabilistic (Semantically secure) encryption and the color features I s used for security by using deterministic encryption on image colour information. This methodology allows privacy preserving CBIR based on colour information to be performed directly on the outsourced servers with high security. Author shows how to design an outsourced image storage, search, and retrieval framework by leveraging IES-CBIR to avoid most heavy computations to be performed by the client. This framework provides increased scalability, performance and lower bandwidth consumption, allowing client applications to be increasingly lightweight and mobile.

The framework is based on two main components first one is an image encryption component, executed on client devices; second one is storage, indexing, and searching component (in the encrypted domain), executed in the outsourcing server (e.g. a cloud provider). IES-CBIR design allows outsourced image repository systems that support content-based image retrieval (CBIR) based on colour features, while protecting the privacy of both image owners and other users issuing queries. Higher computational performance than previous approaches since it securely moves indexing computations to the cloud provider's infrastructure and avoids public-key and homomorphic cryptography. IES-CBIR also minimizes cipher text expansion and consequently bandwidth and outsourced space requirements.

The results show that IES-CBIR is provably secure, allows more efficient operations than existing proposals, both in terms of time and space complexity.

They have proposed a secure framework for the privacy-preserving outsourced storage, search, and retrieval of large-scale, dynamically updated image repositories, where the reduction of client overheads is a major aspect. In the basis of framework is a novel cryptographic scheme, specifically designed for images, named IES-CBIR. Key to its design is the observation that in images, color information can be separated from texture information, enabling the use of different encryption techniques with different properties for each one, and allowing privacy preserving Content-Based Image Retrieval to be performed by third-party, untrusted cloud server.

C. "A privacy-preserving content-based image retrieval method in cloud Environment" Elsevier Inc. 2017

A privacy-preserving content-based image retrieval method based on orthogonal decomposition is proposed in the paper. The image is divided into two different components, for which encryption and feature extraction are executed separately. As a result, cloud server can extract features from an encrypted image directly and compare them with the features of the queried images, so that users can thus obtain the image. Different from other methods, the proposed method has no special requirements to encryption algorithms, which makes it more universal and can be applied in different scenarios. Experimental results prove that the proposed method can achieve better security and better retrieval performance. More effective feature extraction algorithm to improve retrieval accuracy. Multi-party computation and homomorphic encryption based methods are secure, but they require complex computation and communication to be used in low-profile devices and large-scale systems; Distance-preserving randomization methods provide the advantage of high efficient and less user-involvement, but not suitable for high requirements to security, these problems are addressed in this proposed system. By using orthogonal decomposition, image can be divided into encryption field and feature extraction field, therefore encryption operation and feature extraction can be executed separately, which make it possible to let CSPs get image features of encrypted image directly without decrypt it and compare with features of queried image. Final results show the security by orthogonal composition. This method is different from other methods as it does not require cipher algorithms. Result shows that the proposed scheme has good encryption security and better retrieval performance. This is the first time that orthogonal decomposition is used in privacy-preserving image retrieval. In proposed system no need of specific cryptographic algorithm. Author uses Advanced Encryption Standard (AES) to encrypt data. Security against cryptographic attacks should be satisfied. With this method, CSP can retrieve image from encrypted image database directly without violating data privacy. Different from other methods reported in the literatures, the proposed method has no restrictions in using special encryption algorithms, which makes the proposed method more universal, thus can accommodate different kinds of applications. Applying more effective feature extraction algorithm to improve retrieval accuracy is still an open problem. By orthogonal transform, an image is decomposed into components of two orthogonal fields, therefore encryption and feature extraction can be operated separately.

D. "A Content-Based Image Retrieval Scheme Using an Encrypted Difference Histogram in Cloud Computing" mdpi 20117

In this scheme, a secure CBIR scheme based on an encrypted difference histogram (EDH-CBIR) is proposed. Firstly, the image owner calculates the order or disorder difference matrices of RGB components and encrypts them by value replacement and position scrambling. The encrypted images are then uploaded to the cloud server who extracts encrypted difference histograms as image feature vectors. To search similar images, the query image is encrypted by the image users as the image owner does, and the query feature vector is extracted by the cloud server. The Euclidean distance between query feature vector and image feature vector is calculated to measure the similarity. A secure CBIR scheme is proposed by using encrypted difference features. A specially designed image encryption method is proposed to support the feature extraction directly from the cipher text domain. In EDH-CBIR, users only need to complete the work of image encryption; the feature extraction and index establishment will be completed by cloud server, which will largely reduce the user's work. This paper takes the statistical characteristics of difference histogram into account, and considers two difference calculation methods. The retrieval accuracy and security in the two situations are tested and analysed. The scheme encrypts the image by difference matrix calculation, difference value replacement, and difference position scrambling. On the basis of this scheme, they compare it with the ECH-CBIR scheme, the GODH-CBIR scheme, the ODH-CBIR scheme and the DDH-CBIR scheme, and the experiments show that encrypted difference histogram feature has advantages. However, both the EDDH-CBIR and the EODH-CBIR scheme have the problem of security risks under the KBP model. Need more efficient encryption methods to improve the security of the EDH-CBIR scheme.

E. "Privacy-preserving Image processing in the Cloud" IEEE 2018

They explore various image processing tasks, including image feature detection, digital watermarking, and content-based image search. The state-of-the-art techniques, including secure multiparty computation and homomorphic encryption are investigated. A detailed taxonomy of the problem statement and the corresponding solutions is provided. This article studies the problem of high computation complexity. To solve the problem, build a system model and formulating design targets. After that, state-of-the-art techniques are introduced, including homomorphic encryption, secure multiparty computation, and so on. Through the analysis, they find that the balance among design targets: functionality, security, and efficiency makes it difficult to solve the problem by applying only one technique. The integration of different techniques instead of traditional cryptography tools is the most promising research direction in this area. Also, considering the prevalence of JPEG compression among some data, privacy-preserving decompression of JPEG file as a special case of privacy-preserving DCT computation is also a promising research direction in this area.

Table1: Available Approaches.

Sr. No.	Paper Title /Publication	Author	Methodology Used	Performance
1.	A Privacy-preserving and Copy-deterrence Content-based Image Retrieval Scheme in Cloud Computing IEEE Transcation.2016	Zhihua Xia, Xinhui Wang, Liangao Zhang, Zhan Qin, Xingming Sun, and Kui Ren	kNN algorithm, A watermark-based protocol, localitysensitive hashing	Improved search efficiency.
2.	Practical Privacy-Preserving Content-Based Retrieval in Cloud Image Repositories. IEEE 2016	Bernardo Ferreira, Rodrigues, Jo˜ao Leit ˜ao, Henrique Domingos	IES Encryption	High performance and scalability
3.	A privacy-preserving content-based image retrieval method in cloud Environment. Elsevier Inc. 2017	Yanyan Xu , Jiaying Gong , Lizhi Xiong , Zhengquan Xu , Jinwei Wang , Yun-qing Sh	Orthogonal decomposition in cloud environment is proposed in the paper	better security and better retrieval Performance.
4.	A Content-Based Image Retrieval Scheme Using an Encrypted Difference Histogram in Cloud Computing MDPI 2017.	Dandan Liu , Jian Shen , Zhihua Xia ,Xingming Sun	Encrypted difference histogram (EDH-CBIR	Encrypted difference histogram feature has advantages. problem of security risks
5.	Privacy-preserving Image Processing in the Cloud IEEE 2018	Zhan Qin, Jian Weng, Yong Cui, Kui Ren	Homomorphic encryption.	Robust image-processing based applications on devices with limited computation power,

III.PROPOSED SYSTEM

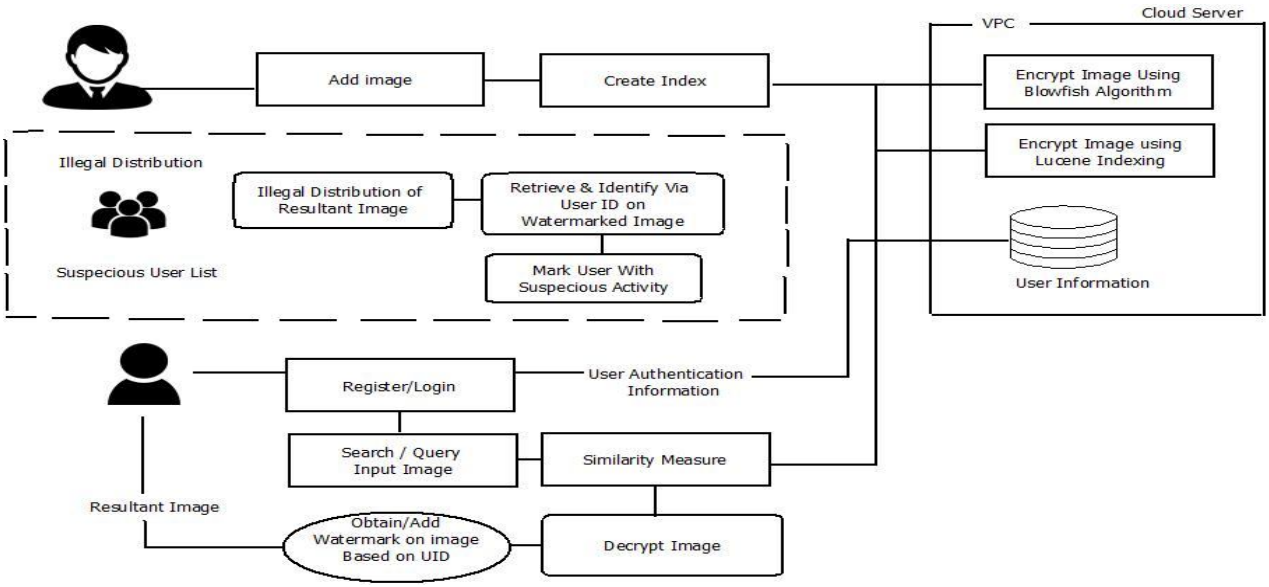


Figure 1 Basic Flow Diagram of Proposed Architecture

A. Image Owner

Image owner outsource their image collection to cloud server in encrypted form having ability to search over encrypted images. Firstly, the image owner extracts the feature vectors from image, then constructs a secure searchable index on feature vector. Then encrypted image collection and index are outsourced to the cloud server.

N images M= {m1,m2,m3,…....m_n}

Encrypted Image C={c1,c2,c3,…...c_n}

Feature Vector F= {f1,f2,f3……..f_n}

Image owner also authorize image user. The image owner sends the authentication information of authorized users to the cloud server who will take the responsibility to verify the identity of user in search requests. In addition, the image owner sends the identities of the authorized users to WCA for watermark generation [1].

Blowfish Encryption algorithm is a symmetric block cipher of block 64 bit used for image encryption. This algorithm has been used because it requires less memory. It uses only simple operations, therefore, it's easy to implement. It's a 64-bit block cipher and is a fast algorithm for encrypting data.

Lucene Indexing

B. Image User

Image users are authorized user who retrieve image from cloud server. To search image, image user needs to generate trapdoor TD for query image and then submit the trapdoor and his identity to cloud server. Image user receive query image, then user decrypt the image by using secrete key shared by image owner.

C. Cloud Server

Cloud server store the collection of encrypted images M and their index I of image owner and process the query request of image user. Cloud user request to cloud server for images, after completion of search operation requested images sent to the user with unique watermark. Server embeds the watermark to retrieved images for supporting copy deterrence. It prevents illegal distribution of images. Illegal distribution of images can be traced by watermark extraction.

IV. CONCLUSION

Content based image retrieval in cloud computing using lucene(Linear) indexing focuses on how the user can effectively retrieve their private documents from cloud while preserving privacy of their documents. Proposed system also provides the watermarking based scheme which provides solution for illegal copy distribution of document.

Many applications are not able to cope with much cost, especially for online applications and mobile devices that use large data sets, such as image data storage with CBIR services. Proposals available in this domain remain in the theoretical domain, as the group requires full homomorphic coding, which is still too expensive. However, some homomorphic encoding schemes and symmetric-solution solutions designed to solve specific search problems. By addressing these challenges, proposed system offers a new security framework for efficient search and retrieval of search, recovery and recovery information.

REFERENCES

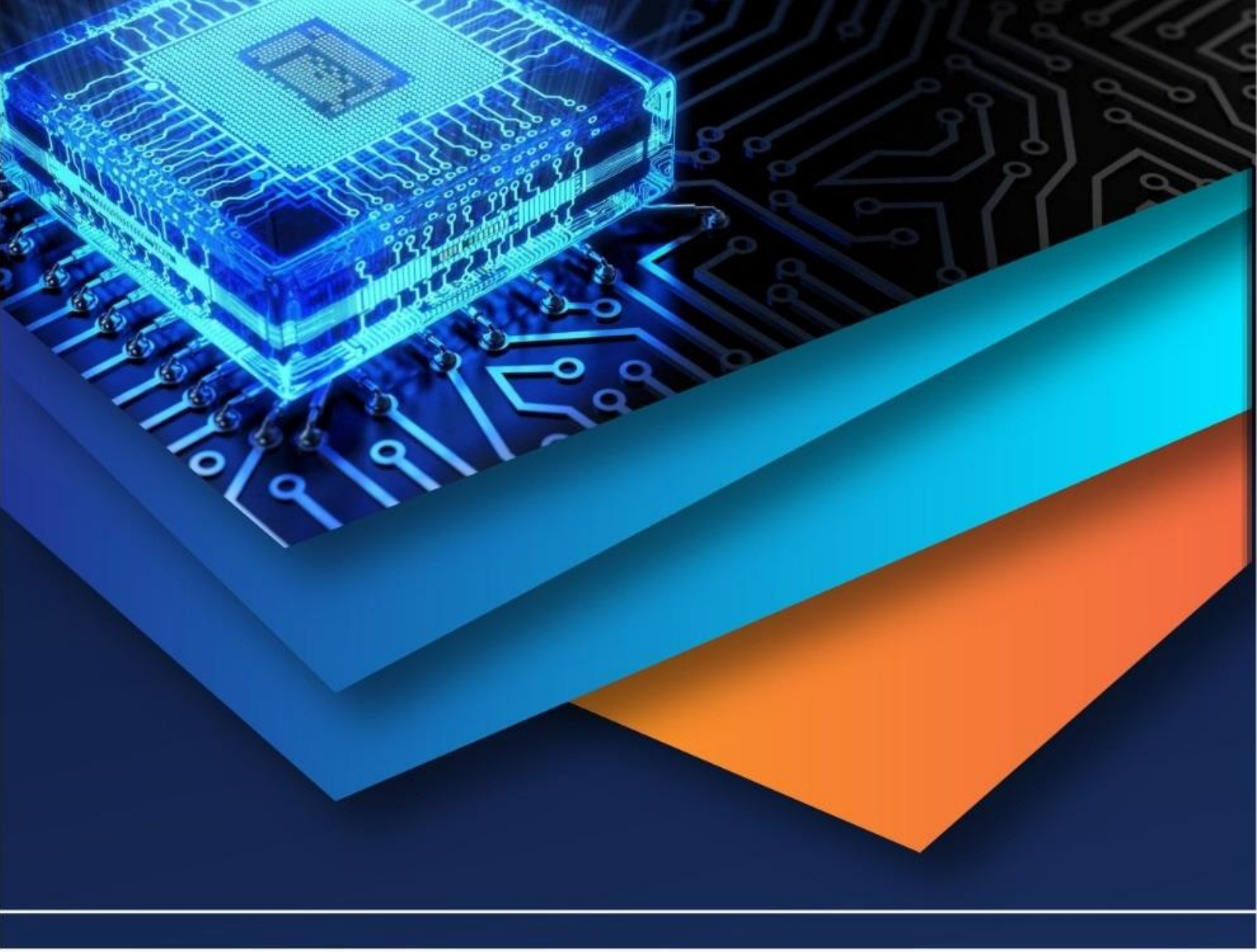
[1] "A Privacy-preserving and Copy-deterrence Content-based Image Retrieval Scheme in Cloud Computing" IEEE Transcation.2016. Zhihua Xia, Xinhui Wang, Liangao Zhang, Zhan Qin, Xingming Sun, and Kui Ren.

[2] "Practical Privacy-Preserving Content-Based Retrieval in Cloud Image Repositories. IEEE 2016" Bernardo Ferreira, Rodrigues, Jo˜ao Leit ˜ao, Henrique Domingos.

[3] "A privacy-preserving content-based image retrieval method in cloud Environment" Elsevier Inc. 2017 Yanyan Xu , Jiaying Gong , Lizhi Xiong , Zhengquan Xu , Jinwei Wang , Yun-qing Sh.

[4] "A Content-Based Image Retrieval Scheme Using an Encrypted Difference Histogram in Cloud Computing " MDPI 2017. Dandan Liu , Jian Shen , Zhihua Xia ,Xingming Sun

[5] "Privacy-preserving Image processing in the Cloud" IEEE 2018. Zhan Qin, Jian Weng, Yong Cui, Kui Ren.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY