SIMILARITY IMAGE RETRIVEAL BASED ON QUERY IMAGE AND DISTANCE FEATURE EXTRACTION CALCULATION METHODS

Dr. Ravindra Krishna Chandar, P. Manoj Prabhakar, S. Kavinraja, S. KavinKumar

Associate Professor, Student,Student,Student
Computer Science And Engineering,
Paavai Engineering College (Autonomous)Namakkal-TamilNadu,

Abstract -Secure picture recovery just on cloud has become increasingly popular in recent years as a result of the introduction of cloud computing. People frequently use the internet to manage and store their sensitive information, the typical method would be to encrypt the information in order to guarantee the secrecy of the data. Proposed technique make use of a balancing index tree. Features from query image was extracted using SIFT (Scale Invariant Feature Transformation) algorithm. Then implement man Hattan distance calculation for detecting suitable index based on query features. Finally, extract the relevant images based on uploaded user-submitted image attributes that have been linked with distance measurements.

IndexTerms–Deep Learning Technique,SHIFT(Scale Invariant Feature Transformation)

Introduction

This project is more than just a digital tool;Similarity-based image search is a valuable technique for finding visually similar images. However, it has limitations related to the accuracy of feature extraction and insensitivity to context. Similarity-based search in the cloud addresses privacy concerns by encrypting images before storage. Images are converted into feature vectors, encrypted, and stored. When a user queries for an image, the encrypted query is compared with stored encrypted feature vectors using a privacy-preserving similarity metric. The cloud server returns encrypted feature vectors of similar images, which are decrypted client-side. This technique ensures privacy while enabling efficient image retrieval, beneficial in sensitive domains like medical imaging or national security.

I. Objectives

•The objective of this project is to efficiently and effectively find and present images from a database that are visually similar to a given query image.

• This concept allows users to search for images based on their visual content, making it possible to retrieve relevant images even when there are no textual descriptions or tags associated with the images in the database.

II. PROBLEM DEFINITION

Content based image retrieval is well known technology being used for the retrieval of images from large database. This image retrieval is a challenging topic that has been a research focus from many years. This has proven very much important because of its applications like face recognition, fingerprint recognition, pattern matching, verification and validation of images. The image retrieval is also called image classification in large database systems. In the past few years, there has been tremendous growth in database technology to store and retrieve large number of images. This requirement creates a demand for software systems for effective fast image retrieval from large database systems .The demand and use of multimedia applications in present world creates the need of content based image search and retrieval.

III. RELATED WORK

Gao, Peng,….[1] Propose a secure and efficient scheme to outsource the item reputation on hyperspectral remote sensing images to the untrustworthy cloud server. The proposed scheme can defend the privacy of the computation input and output. Also, here broaden an effective verification approach in proposed scheme that could stumble on the misbehavior of cloud server with the choicest chance. Here advocate the primary comfortable outsourcing scheme for item popularity on hyperspectral remote sensing images, which lets in the resources-limited devices to perform object reputation in a secure and efficient way. Particularly, the performance of the image recognition is stepped forward via delegating the workload to a cloud server, at the same time as the privacy of inputs and outputs are ensured. To make certain the privateness of enter and output, right here advocate a singular essential transformation-based totally approach. Also, assemble the verification method so that the patron can discover any misbehaviour of the cloud server. And the verification approach will now not boom the computational burden of the patron. Finally compare proposed scheme thru substantially theoretical analysis and experiments. Theoretically analyze the correctness, safety and performance of the proposed scheme. The proposed scheme utilizes elementary rotation matrices to protect the privacy of the computation input and output. The customer in the proposed scheme can verify the correctness of the lower back consequences from the cloud server.

Shen Meng,..[2] Implement a privacy protection CBIR scheme that helps Multiple Image owners with Privacy Protection (MIPP). Here encrypt image capabilities with a at ease multi-party computation technique, which allows picture owners to encrypt photograph features with their own keys. This permits efficient photo retrieval over features gathered from multiple sources, while guaranteeing that image privateness of an individual image owner will not be leaked to different image proprietors. Also suggest a brand new technique for similarity dimension of pictures that could avoid revealing picture similarity statistics to the cloud. MIPP operates in the identical way as existing schemes in the first class, which outsources encrypted photos in conjunction with their encrypted image functions to the cloud. In order to cope with the demanding situations of supporting a couple of photo proprietors, in this approach first encrypt photographs with a key circulate and encrypt the corresponding photo functions by means of the comfy multi-birthday celebration computation technique, and then advise a unique technique to degree the picture similarity; this will assist to avoid revealing the image similarity records in cloud to a positive extent. In the proposed MIPP, multiple picture owners are allowed to encrypt photographs and picture features by their precise secret image encryption keys. This enables efficient picture retrieval over photos accrued from multiple sources, whilst presenting guarantees that image privateness of a character photo owner will not be leaked to other image owners. Thus, the proposed MIPP can meet the practical necessities in actual-global packages.

Wu Tong,…[3] Implement an RDIC scheme to concurrently verifying the uploaded data content and duration of data storage represented via an updatable timestamp through the third party auditor (TPA). Also, proposed scheme achieves indistinguishable privacy (IND-privacy) in the direction of TPA for each facts content and timestamp. To bind the data content and timestamp inside the authenticator and guide efficient timestamp replace, here assemble the authenticator with the randomizable shape-maintaining signature (SPS). Additionally, this approach utilizes the Groth-Sahai evidence and variety evidence to offer the IND-privacy and guarantee the timestamp validation within the auditing section. To guide the above PAYG pricing version, a entire-fledged RDIC protocol wishes to offer the records integrity and timestamp validation simultaneously. Moreover, it need to additionally assist efficient renewal of the garage issuer through updating the timestamp. It is not trivial to assemble an RDIC scheme helping timestamp validation and update by way of right now enhancing the existing RDIC schemes. The important solution is that the CSP sends the timestamp to the third party verifier for verification, which, but, has the issues that the timestamp ought to be discovered to the verifier and isn't suitable inside the third party auditing placing wherein the TPA want to now not look at any information about the data storage alongside the timestamp if user wants to keep away from any inference assaults that can be accomplished with the resource of the TPA to deduce the information content cloth. Moreover, a way to allow the timestamp to be effectively updated if user desires to allow customers to resume their garage duration is a difficult hassle. The trivial answer is to compute new authenticators with the brand new timestamp. However, the client has to download all the data with a view to update the authenticators, which is impractical. Also, to lessen the computation burden at the consumer's aspect, it is right to allow the CSP to replace the authenticators for the brand new timestamp.

Song ,Weiwei's,…[4] Propose a unique deep hashing convolutional neural network (DHCNN) to concurrently retrieve the same pictures and classify their semantic labels in a unified framework. In greater element, a convolutional neural community (CNN) is used to extract high dimensional deep capabilities. Then, a hash layer is perfectly inserted into the community to switch the deep features into compact hash codes. In addition, a fully connected layer with a softmax characteristic is performed on hash layer to generate class distribution. Finally, a loss characteristic is elaborately designed to concurrently remember the label loss of each picture and similarity lack of pairs of features. In extra element, here first undertake a CNN to extract high-dimensional deep functions from raw remote sensing pictures. Then, a hash layer is flawlessly inserted into the CNN to encode the high-dimensional deep functions to lowdimensional hash codes. In addition, a fully linked layer with a softmax function is done on hash layer to generate class distribution. Finally, here elaborately layout a loss characteristic to teach DHCNN, wherein the label data of each photo and similarity facts of pairs of snap shots are simultaneously considered to improve the capability of representation of functions. Once DHCNN is educated enough, for a question picture, can generate its hash code with the aid of binarizing the output of hash layer, then, the retrieval may be without difficulty finished thru Hamming distance ranking. In addition, the semantic labels of snap shots, inclusive of the question photograph and its comparable pictures, can be obtained by using feeding their semantic features into the softmax classifier.

IV. PROPOSED WORK

As Cloud Computing grows, sensitive information like emails, health records, and government documents are centralized in the cloud. Storing data there relieves owners of storage and maintenance burdens, offering high-quality storage services on demand. Sharing data with many users is common, where users may want to selectively retrieve specific files through image-based search instead of retrieving all details, which is impractical. The SIFT algorithm is vital, detecting and describing invariant keypoints in images, capturing unique characteristics regardless of scale, rotation, or illumination changes. Extracting SIFT features allows robust matching and retrieval of similar images, aided by the Manhattan distance (L1 distance), measuring similarity between descriptors by computing absolute differences. Comparing Manhattan distances between query image descriptors and those in the database ranks and retrieves images with similar visual content efficiently, offering robustness and effective retrieval capabilities.

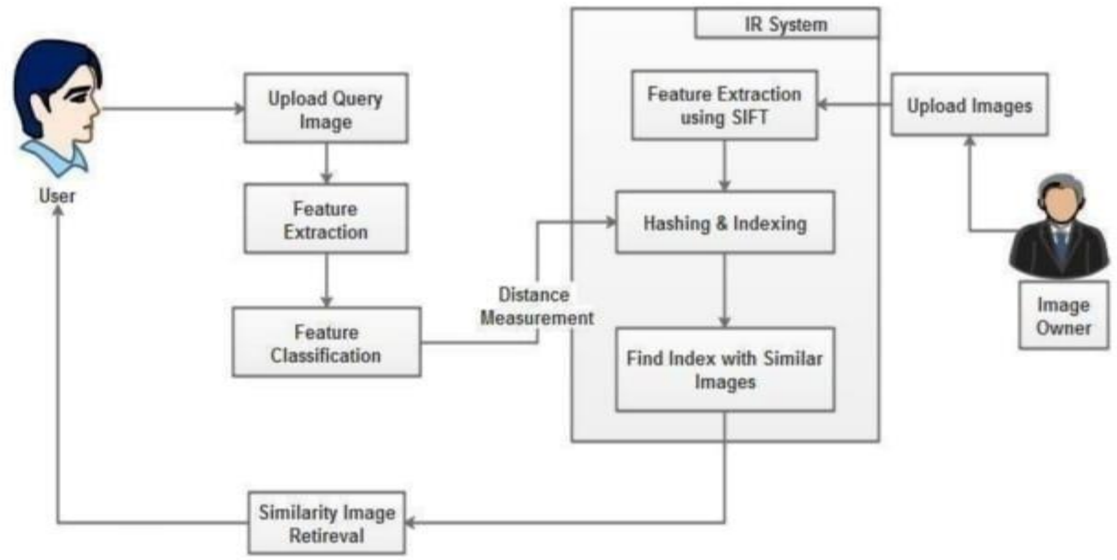

Fig 1: Proposed Architecture

QUERY IMAGE PROCESSING:

Image processing is a technique that involves analysing and manipulating images to extract useful information or enhance their visual quality. In the context of retrieving similarity-based images in a cloud module, image processing plays a crucial role in comparing and matching images based on their visual characteristics. In a cloud module designed for similarity-based 16 image retrieval, the image processing pipeline typically consists of several stages. The input images are pre-processed to remove noise, correct for lighting conditions, and normalize their size and orientation. This step ensures that the images are in a consistent format for further analysis.

FEATURE EXTRACTION

Feature extraction refers to the process of transforming raw data into numerical features that can be processed while preserving the information in the original data set. It yields better results than applying machine learning directly to the raw data. Feature extraction can be accomplished manually or automatically:

Manual feature extraction requires identifying and describing the features that are relevant for a given problem and implementing a way to extract those features. In many situations, having a good understanding of the background or domain can help make decisions. Over decades of research, engineers and scientists have developed feature extraction methods for images, signals, and text. An example of a simple feature is the mean of a window in a signal. Automated feature extraction uses specialized algorithms or deep networks to extract features automatically from signals or images without the need for human intervention. This technique can be very useful when user want to move quickly from raw data to developing SIFT algorithms.
Step 1: Pre-processing: In pre-processing grayscale conversion, invertion and smoothing of image is done using Filter method.
Step 2: Shape Extraction: Shape features are extracted based on x and y co-ordinate representation of image using SIFTS algorithm.
Step 3: Color Extraction: Histogram based R, G, B Color Coherence Vector are analyzed.

INDEX DETECTION

In the index finding stage of the similarity image retrieval process, the Manhattan distance calculation plays a crucial role in comparing the feature vectors of query images with the stored feature vectors in the index. After the feature extraction step, each image in the database is represented by a feature vector, which captures the relevant visual characteristics of the image. The feature vectors may contain elements corresponding to colour histograms, texture descriptors, or other relevant image attributes. To find images that are similar to a query image, the Manhattan distance is computed between the query image's feature vector and the feature 17 vectors of the stored images. The Manhattan distance measures the total "distance" between two vectors by summing the absolute differences between their corresponding elements. This distance metric is suitable for similarity comparisons because it considers both positive and negative differences between elements.

SIMILARITY IMAGE RETRIEVAL

The system can rank the stored images based on their similarity to the query image by computing the Manhattan distance for all feature vectors in the index. Images with smaller Manhattan distances are considered more similar to the query image, while those with larger distances are considered less similar. The Manhattan distance calculation provides a quantitative measure of the dissimilarity between images and enables the retrieval system to efficiently find images that are visually similar to the query image. By incorporating the Manhattan distance calculation into the index finding stage, the system can effectively identify and present the most relevant images to the user, enhancing the overall image retrieval experience.

V. CONCLUSIONS

Similarity-based image retrieval in the cloud using query image processing, SIFT feature extraction, and Manhattan distance calculation provides an effective way to search and retrieve visually similar images. The combination of these techniques allows for efficient indexing and retrieval, making it possible to handle large-scale image databases. By leveraging cloud computing resources, the system can scale to accommodate a growing number of images and users. As image collections continue to expand, similarity-based image retrieval in the cloud becomes an invaluable tool for various applications, including image search engines, content-based image retrieval systems, and visual recommendation systems.

VI. REFERENCES

[1] Gao, Peng, Hanlin Zhang, Jia Yu, Jie Lin, Xiaopeng Wang, Ming Yang, and Fanyu Kong. "Secure cloud-aided object recognition on hyperspectral remote sensing images." IEEE Internet of Things Journal 8, no. 5 (2020): 3287-3299.
[2] Shen, Meng, Guohua Cheng, Liehuang Zhu, Xiaojiang Du, and Jiankun Hu. "Contentbased multi-source encrypted image retrieval in clouds with privacy preservation." Future Generation Computer Systems 109 (2020): 621-632.
[3] Wu, Tong, Guomin Yang, Yi Mu, Rongmao Chen, and Shengmin Xu. "Privacy-enhanced remote data integrity checking with updatable timestamp." Information Sciences 527 (2020): 210-226.
[4] Fu, Anmin, Shui Yu, Yuqing Zhang, Huaqun Wang, and Chanying Huang. "NPP: A new privacy-aware public auditing scheme for cloud data sharing with group users." IEEE Transactions on Big Data 8, no. 1 (2017): 14-24.
[5] Song, Weiwei, Shutao Li, and Jón Atli Benediktsson. "Deep hashing learning for visual and semantic retrieval of remote sensing images." IEEE Transactions on Geoscience and Remote Sensing 59, no. 11 (2020): 9661-9672.
[6] Ravishankar, B., Prateek Kulkarni, and M. V. Vishnudas. "Blockchain-based database to ensure data integrity in cloud computing environments." In 2020 International Conference on Mainstreaming Block Chain Implementation (ICOMBI), pp. 1-4. IEEE, 2020.
[7] Liu, Yishu, Conghui Chen, Zhengzhuo Han, Liwang Ding, and Yingbin Liu. "Highresolution remote sensing image retrieval based on classification-similarity networks and double fusion." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13 (2020): 1119-1133.
[8] Shao, Zhenfeng, Weixun Zhou, Xueqing Deng, Maoding Zhang, and Qimin Cheng. "Multilabel remote sensing image retrieval based on fully convolutional network." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13 (2020): 318- 328.