جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة ديالى
كلية العلوم
قسم علوم الحاسبات

استرجاع الصور بشكل آمن على اساس الشبكة العصبية التلافيفية

رسالة مقدمة
الى كلية العلوم في جامعة ديالى وهي جزء من متطلبات نيل
شهادة الماجستير في علوم الحاسبات

من قبل

شيماء متعب سعدون

بإشراف

أ.د زياد طارق مصطفى الطائي

1445 هـ

2023 م

Chapter One

General Introduction

Chapter One

General Introduction

1.1 Overview

With the rapid development of multimedia and the internet, vast amounts of images are being produced and disseminated. The question of how to efficiently store and communicate such large amounts of data has become an important issue. Outsourcing images to the cloud is a natural solution, because of its many benefits, fast flexibility of resources, pricing, and management based on use, network access from anywhere, shared resources, and self-service on request are all examples of these trends, Image retrieval is another service that may be performed in the cloud [1,2]. A wide variety of areas can use this technology, including the visual arts, multimedia, satellite image retrieval, internet commerce, medical imaging, forensic science, and the news media. These are just a few of the many fields that can benefit from it, Simultaneous localization and mapping (SLAM) rely heavily on image retrieval, which has found widespread use in applications like mobile robotics and autonomous driving[3,4].

Image retrieval relies heavily on Content-Based Image Retrieval (CBIR) techniques, the primary purpose of which is to discover and present all images that share visual content with a specified query image. Data owners frequently employ CBIR techniques to improve retrieval accuracy and efficiency[5, 6], and cloud image storage allows authorized users to access images via query easily. While cloud computing provides many benefits, it is essential to remember that not all data owners are comfortable entrusting their information to a remote server[7].The image outsourcing process might result in a variety of breaches of privacy. The images

always contain a wealth of sensitive information, such as personal data or clinical facts about the patient. As a result, it is necessary to create a CBIR system that protects users' privacy when using cloud computing[8].

Sensitive images should be encrypted before being transferred to the cloud to safeguard users' privacy and promote data security; however, the encryption procedure might impede several typical functions in the cloud environment, such as image retrieval. One of the most promising approaches to image retrieval is content-based image search (CBIR). Image feature extraction and feature distance comparison are their defining characteristics. However, after the picture has been encrypted, it becomes impossible to employ the CBIR approach because of the unpredictability introduced by the encryption process[9].

Several strategies have been developed to challenge secure retrieval technology in recent years. These strategies vary in how they encrypt images and extract features from encrypted images. There are two issues with the approaches that have been suggested. First, security and efficiency are sometimes at odds with one another. Lightweight encryption techniques, such as permutation and substitution, are fast but insecure, whereas multiparty computing and homomorphic encryption algorithms are secure but impractical due to their high computation costs. Second, efficiency and precision in retrieval are competing goals. Although several approaches employ low-level data like color, texture, form, etc. for image retrieval, the retrieval accuracy seldom meets the requirements of practical applications due to the "semantic gap" between visual cues and the richness of human semantics.

1.2 Problem Statement

With the increase in the number of Internet users and the advent of the cloud concept, in addition to the global reliance on cloud storage, the urgent need for secure image recovery has emerged. So, secure and correct retrieval of images from the billions of images on the cloud is the problem of this thesis.

1.3 Related Work

Previous work that is relevant to this thesis is presented in this section:

- **In 2016, Xia et al.** [10]**,** They proposed CBIR on encrypted images without giving the cloud service provider (CSP) private data. Image representation requires feature vector extraction. This approach uses MPEG-7's four visual descriptors: Color Layout Descriptor(CLD), Scalable Color Descriptor (SCD),Edge Histogram Descriptor (EHD),and Color Structure Descriptor (CSD).Secure K-Nearest Neighbor(K-NN) encrypts feature vectors. To speed up the search, they created pre-filter tables using Locality-Sensitive Hashing (LSH) and encrypted picture pixels using a typical stream cipher. They also used watermarks to prevent authorized inquiry users from duplicating and sharing images with unauthorized third parties. Using this watermark-based protocol, the cloud serviceadded a unique watermark onto each encoded image before transmitting it to the querying user. Thus, the extraction of watermark could identify the unauthorized query user who posted the photograph. On the Corel10K dataset, they found AP@K for CSD =20%, k=100,reduced

precision enhanced search efficiency, Average search times =13.8144ms, with PSNR=36.02dB and SSIM=0.8531.

- **In 2017, Xia et al.** [11]**,** They developed a cloud-based, content based image retrieval method that safeguarded individual privacy. Protecting the feature vectors involved in MPEG-7, two visual characterizations were defined as standards. (CLD) and (EHD) were used in the scheme. The vectors of feature were encrypted using the secure KNN method, allowing the cloud server to rank the search results quickly without incurring extra communication costs. They used (LSH) to build the pre-filter tables for similar group photos. They created a two-level index. The higher level stored the pre-filter tables, enhancing the search's effectiveness. The lower level was the one-to-one map index that could be used to sort search results. The Corel10K dataset was used to conduct the experiments. CLD had an AP@K=10% whereas EHD's was 10.55%, k=100. These accuracies were reduced in exchange for a faster search time. Consumption time decreased by 67.12% on average for CLD and 58.86% for EHD.
- **In 2017, Xu et al.** [12]**,**In this study, a proposal was made for a content-based image retrieval approach that would protect users' privacy and would be based on orthogonal decomposition in a cloud context. An image was split into components that belong to two orthogonal fields using an orthogonal transform. As a result, encryption and feature extraction was performed independently. Coefficients of the Discrete Cosine Transform(DCT) were used to extract features. After orthogonal composition, the final data were formed by integrating two different components that are independent of one another. Using this technology, (CSP) can obtain images

straight from an encrypted picture database without infringing on the users' privacy rights. The solution that is being proposed does not place any limitations on the use of specialized encryption methods. Experiments demonstrated on the Corel 1k dataset, the Corel 10k dataset, and the Inria Holidays dataset that the method attained Precision= 0.55, Recall=0.1 , mAP= 46.62%, PSNR equal to 8.9547 dB.

- **In 2019, Qin et al.** [13]**,** Deep learning and adaptive weighted fusion are the two technologies utilized in the image retrieval approach suggested in this research article. To begin, the process of extracting low-level features such as Bag of Words (BOW), (EHD), and (CLD), as well as high-level semantic features (CNN) of pictures. Second, to reduce the size of a high-level semantic feature that had 1024 dimensions, a Principle Component Analysis (PCA) was performed, and then each of the three features was binarized. After then, elements of this type are adaptively fused. The final step is constructing a prefilter table for fusion features to enhance the search's efficiency using (LSH) technique. In order to safeguard the confidentiality of the fused features and pictures, (KNN) technique and the logistic encryption approach were utilized. Experiments demonstrated on the Corel10k dataset that the approach obtained precision@k= 46.9%, k=100, retrieval time = 7856 ms.
- **In 2020, Gu et al.** [14]**,** In this research, the authors examined the challenges associated with Multi-Source Privacy-Preserving Image Retrieval (MSPPIR). They suggested a system that used a unique JPEG image encryption scheme that is designed for Multi-Source content-based image Retrieval. This system could satisfy the needs of MSPPIR, such as the secure retrieval of information in continual

rounds from many sources and the union of different sources for improved retrieval services. They created a strategy that depended on randomization encryption, the bit XOR and the permutation were utilized, and the (BOW) model was applied to derive characteristics from encrypted images. Experimental results were obtained on Corel1K and Corel10K, with a precision@k= 17.79%, k=100.

- **In 2020, Shen et al.**[15]**,** This study suggested a secure CBIR solution using the Multiple Image Owners with Privacy Protection (MIPP) approach. After Edge Histogram Descriptors (EHD) were used to extract imagee features, a stream cipher was applied to encrypt both the images and the features retrieved from them. According to the Corel 10K dataset, the retrieval score is comparable to the usual Euclidean distance criteria, with a precision@k=22.64%, k=100, and an average search time=50ms.
- **In 2021, Pan et al.**[16]**,** this study suggested an enhanced CNN-based hashing technique for retrieving encrypted images. In order to enhance the CNN's capacity for representation, the image size was first raised, and to lower the parameters and computational cost of the CNN, a lightweight module was introduced to replace parts of the modules. A hash layer was finally implemented to create a small binary hash code. The hash code was employed throughout the retrieval procedure for encrypted images; experimental results on the Corel10K dataset were obtained with precision@k= 69.91%, when k=100, retrieval time=1.249s.
- **In 2021, Punithavathi et al.** [8]**,** This paper presents a Secure Image Retrieval System employing Inception and ResNet v2 (SIRS-IR) and Multiple Share Creation (MSC). Inception and ResNet v2 model-based feature

extraction are proposed. The MSC procedure generates multiple shares, and the Double Chaotic Logistic Map (DLCM) approach encrypts them. The cloud server stores encrypted shares and feature vectors with the image identification number. On the Corel10K dataset, the SIRS-IR system had Avg. Precision=0.94, Avg. Recall=0.89, MSE=0.0652, PSNR=60.51 dB.

- **In 2022, Ma et al.**[17]**,** They presented an image retrieval system that protects users' privacy based on CNN features. This system would consider the sensitive nature of data during the preprocessing, storage, and search stages of cloud computing. Specifically, they developed a hybrid encryption method that includes Channel Encryption, Sequence Encryption, and Position Encryption. This method can safeguard photos' color and texture information, which is necessary for preventing unauthorized cloud servers from revealing critical data. Meanwhile, the cloud server can extract semantic features from encrypted photos using an upgraded DenseNet-121 model version. It can then do feature similarity matching to provide all retrieval results comparable to those extracted features. On Holidays and Corel10K datasets the suggested system achieved PSNR = 31.98dB,AP@K=44.06% , when k=100=44.06% when k=100, and Average search time = 34.36ms.

Related Works Summary Table (1.1).

Table 1.1: Related Works Summary.

No.	Study	Year	Dataset	Descriptor	Result
1	**Xia et al.** [10]	2016	Corel10k	CSD, CLD, EHD, and SCD	Ap@100=20 %, Average search time =13.8144ms, PSNR= 36.02dB, SSIM=0.8531

2	**Xia et al.** [11]	2017	Corel10k	CLD, EHD	CLDAp@100=10% EHD Ap@100= 10.55 %
3	**Xu et al.** [12]	2017	Corel1k Corel10k Holidays	DCT	Precision= 0.55 Recall=0.1 , mAP= 46.62%, PSNR = 8.9547dB
4	**Qin et al.**[13]	2019	Corel10k	CNN, BOW , CLD and EHD feature	precision@100= 46.9%, retrieval time = 7856 ms
5	**Gu et al.** [14]	2020	Corel1k Corel10k	BOW	precision@100= 17.79%.
6	**Shen et al.**[15]	2020	Corel10k	EHD	AP@100=22.64%,average search time=50ms, Time consumption of index construction=419.724s
7	**Pan et al.**[16]	2021	Corel10k	CNN	precision@k= 69.91%, when k=100. Time consumption of retrieval=1.249s Parameters=176.64 MFLOPs=158.25 Feature extraction = 128.25 s
8	**Punithavathi et al.** [8]	2021	Corel10k	CNN	Avg. Precision=0.94 , Avg. Recall=0.89, MSE=0.0652, PSNR=60.51 dB
9	**Ma et al.**[17]	2022	Corel10k , Holidays	CNN	AP@K=44.06% , when k=100 PSNR = 31.98dB, Feature extraction = 67.2s Average search time = 34.26ms Parameters= 176.54M MFLOPs= 158.21

1.4 Aim of the Thesis

The fundamental aim of this work is to design and implement a secured image retrieval in the cloud. This model is aimed at achieving both security and accuracy in its image retrieval process. Accordingly, the following objectives have been realized:

1. The model's design relies on utilizing the capabilities of a deep convolutional network, which is known for its effectiveness in image-related tasks.

2. To enhance the security of the retrieved images, the thesis incorporates a visual cryptography technique. This technique is specifically based on the Lotka-Volterra Chaos Map, suggesting a new approach to image security.

3. an optimized DenseNet-121 model plays a crucial role. It serves as a calibration component, likely used as a feature extractor to enhance the image retrieval process, potentially improving the model's overall performance.

1.5 Contribution

The most important contribution is that the proposed model guarantees image retrieval with stronger security than the general encryption technology, by using visual encryption technology.

1.6 Outline of the Thesis

In addition to the first chapter, which offers a concise overview of secured image retrieval, a statement outlining the problem, recent research on image

retrieval, and the core objective of the thesis is included; the subsequent content comprises a succinct synopsis of the topics addressed in the remaining chapters:

Chapter Two:" Theoretical Background "

It explains the theoretical background of the secure image retrieval approach, visual cryptography and how it works, extracting encrypted features using traditional and modern methods, and calculating image retrieval using similarity measures.

Chapter Three: "Design and Implementation of Proposed Model"

The proposed approach's methodology is declared, which encompasses the stage and algorithms incorporated in the process of retrieving secured images.

Chapter Four: "Results and Discussion"

This chapter gives the experimental results obtained from the Implementation of a secured image retrieval approach.

Chapter Five: "Conclusions and Suggestions for Future Work."

This chapter includes conclusions and future work for the development of a secured image retrieval approach with a list of a number of suggestions for future studies.

الخـــلاصة

لقد حظي مفهوم استرجاع الصور المعتمد على المحتوى (CBIR) باهتمام كبير مؤخرًا نظرًا للأهمية المتزايدة للصور في حياة الناس اليومية. تتطلب الصور مساحة تخزين أكبر مقارنة بالمستندات النصية. ومع ذلك، يجب تشفير الصور المهمة مثل الصور الطبية والشخصية. وهذا يجعل تقنيات استرجاع الصور المعتمد على المحتوى التي تعمل في مجال النص العادي عديمة الفائدة. ولذلك، فإن التحدي المتمثل في تصميم طرق التشفير ذات الأمان العالي واستخراج الميزات القيمة من صور النص المشفر لا يزال يشكل عائقًا أمام استرجاع الصور المستند إلى المحتوى والحفاظ على الخصوصية.

تقدم هذه الأطروحة آلية للتعامل مع "استرجاع الصور الآمن بالاعتماد على الشبكة العصبية التلافيفية" كتطوير واستخدام جديد لاسترجاع الصور. يتكون النموذج المقترح من ثلاث مراحل، سنوضح كل منها بإيجاز فيما يلي: الأولى هي مرحلة تشفير الصور. يتم استخدام التشفير المرئي (VC) استنادًا إلى خريطة الفوضى Lotka_Volterra ثلاثية الأبعاد كخوارزمية لتشفير الصور. تستخدم الخطوة التالية خوارزميات التعلم العميق لاستخراج الميزات من الصور المشفرة. المرحلة الثالثة هي الاسترجاع، والتي تقدم مجموعة فرعية من الصور بناءً على التشابه بين الصور المحسوبة باستخدام ناقل الميزات المسترد من كل صورة.

أظهرت التجارب على مجموعة بيانات Corel10k فعالية النموذج المقترح من حيث الأمان والدقة وكفاءة البحث مقارنة بالأعمال السابقة، حيث حققت mAP@Top-k= 43.8856479، وزمن البحث = s1.00، وMSE =12764.27263، PSNR. = 7.300480125 ديسيبل بين الصورة الأصلية والمشفرة. الصورة التي تم فك تشفيرها مطابقة تمامًا للصورة الأصلية.