Balancing Privacy and Performance: Exploring Encryption and Quantization in Content-Based Image Retrieval Systems

Mohamed Jafar Sadik, Dr. Noor Azah Samsudin, Dr. Ezak Fadzrin Bin Ahmad
Faculty of Computer Science and Information Technology (FSKTM),
Universiti Tun Hussein Onn Malaysia (UTHM), Batu Pahat, Johor 86400, Malaysia

***Abstract*—This paper presents three significant contributions to the field of privacy-preserving Content-Based Image Retrieval (CBIR) systems for medical imaging. First, we introduce a novel framework that integrates VGG-16 Convolutional Neural Network with a multi-tiered encryption scheme specifically designed for medical image security. Second, we propose an innovative approach to model optimization through three distinct quantization methods (max, 99% percentile, and KL divergence), which significantly reduces computational overhead while maintaining retrieval accuracy. Third, we provide comprehensive empirical evidence demonstrating the framework's effectiveness across multiple medical imaging modalities, achieving 94.6% accuracy with 99% percentile quantization while maintaining privacy through encryption. Our experimental results, conducted on a dataset of 1,200 medical images across three anatomical categories (lung, brain, and bone), show that our approach successfully balances the competing demands of privacy preservation, computational efficiency, and retrieval accuracy. This work represents a significant advancement in making secure CBIR systems practically deployable in resource-constrained healthcare environments.**

Keywords*—*Content-Based Image Retrieval (CBIR); Convolutional Neural Networks (CNN): Encrypted data; Feature extraction; Fully Homomorphic Encryption (FHE); medical imaging; privacy; quantization; retrieval accuracy

I. INTRODUCTION

Content-Based Image Retrieval (CBIR) systems play a vital role in managing digital images, as they support the sorting and retrieval of images based on their specific content like colour patterns, texture elements or shape structure beyond the external tags or descriptions. By eliminating manual annotation individuality and its limitations, this approach significantly outperforms traditional keyword-based retrieval methods by automating the task. By extracting features and indexing images according to those extracted data, CBIR systems can deliver more intuitive and reliable outcomes by predicating directly from the visual content of an image [1–3].

Incorporating encryption into CBIR systems introduces Encryption is being considered as an essential security element for CBIR systems because of the higher and severe requirements with a larger number of digital images to be protected from unauthorized access. By using encryption, sensitive information within images stays safe from unauthorized access, while still enabling the CBIR system to retrieve the required images. This is especially relevant in contexts where privacy is paramount and the image data is too sensitive to leave unencrypted [4–8].

The medical sector stands as a prominent example where the integration of encryption with CBIR is not just beneficial but essential. Medical images contain private patient information, and their retrieval requires the utmost care to maintain confidentiality. As such, medical image retrieval systems must adapt to perform effectively on encrypted data, ensuring that patient privacy is maintained without compromising the diagnostic value of the images [6, 9,10].

Investigating the technical side, the effectiveness of CBIR can be enhanced with the use of advanced machine learning models such as transformers and Convolutional Neural Networks. These pre-trained models have shown exceptional ability in feature recognition and extraction from large datasets, making them irreplaceable in encrypted image analysis. Transformers handle data sequence well, which can be critical in understanding image context, while Convolutional Neural Networks excel in pattern recognition due to their layered architectural design [11–16].

Quantization further refines the functionality of Convolutional Neural Networks in CBIR by compressing the models without significant loss of performance. This process reduces the demand on computational resources and is particularly advantageous for deploying sophisticated CBIR in environments with hardware limitations or where swift image retrieval is needed [17–20].

Despite the significant advances in Content-Based Image Retrieval systems, particularly in the realm of medical imagery, there remains a critical gap in the development of a generalized framework that seamlessly integrates robust encryption protocols. Current CBIR systems often face a trade-off between encryption strength and retrieval accuracy, with many falling short in providing a secure yet efficient retrieval process that caters to the highly sensitive nature of medical data [15,19, 21–23]. Additionally, the application of such systems is frequently constrained by the computational power required for processing and retrieving high-resolution medical images, which further complicates their deployment in resource-limited settings commonly found in healthcare environments.

Addressing these challenges, our study aims to propose a comprehensive framework for CBIR tailored to the medical domain. The goal is to craft a lightweight, secure encryption

algorithm that adequately protects patient confidentiality without affecting the system's ability to accurately retrieve and analyze medical images. Encryption alone, however, is not the primary focusing point of this framework. The proposed system also seeks to harness the power of Convolutional Neural Network-based pre-trained models, particularly the VGG-16 architecture well-known for its ability in image recognition tasks, to facilitate the feature extraction process in encrypted domains.

Furthermore, to tackle the issue of computational Furthermore, to tackle the issue of computational efficiency, the integration of a quantization approach is crucial. In order to deploy the vgg-16 model into practical usage, it is necessary that its memory footprint and computational needs are significantly reduced thereby making the solution suitable to be deployed even in work environment having low check for computational power. Our use case is under the same limit, by processing through quantization we hope to not only improve efficiency of our system but still keep a high retrieval performance which has been hard trade-off in many existing CBIR systems.

II. Contributions

There are three major contributions in this work to-wards medical CBIR:

A. Developing an Efficiency-Aware Framework

We illustrate a CBIR system for medical images, focusing on computational efficiency and accuracy as the main goals. The system is built for high-throughput needs of medical diagnostics with the modest computational resources common in healthcare settings.

B. Pre-Trained CNN Model

The cornerstone of our retrieval pipeline around which all other piece's center, the VGG-16 model is a pre-trained Convolutional Neural Network we use for feature extraction. In using this well-developed model, our framework can take advantage of its deep learning power to improve the accuracy in medical image analysis.

C. Advanced Approach to Quantization

The framework uses a variety of strategies to quantization, including max quantization, 99% percentile quantization, and Kullback-Leibler Divergence (KL) based quantization. These methods reduce the model's memory footprint and computational load, which is crucial for scalable and efficient CBIR deployment.

III. Literature Survey

A. Feature Extraction and Encryption Methods for CBIR

In the rapidly evolving field of cloud-based privacy-preserving image retrieval, researchers have embarked on a journey to balance the confidentiality of encrypted images with the necessity of efficient retrieval mechanisms. This literature survey begins with foundational encryption techniques that pave the way for more advanced retrieval methods, highlighting the journey from basic encryption to sophisticated integrations of cryptographic algorithms and machine learning models. It provides a comprehensive over-view of the state-of-the-art methods developed to ad-dress the dual challenges of maintaining privacy and ensuring practical utility in cloud environments.

Starting with encryption strategies for image security, works such as [5, 24, 25] introduce novel schemes focusing on the variation of Discrete Cosine Transform (DCT) coefficients, encryption of DC and AC coefficients using stream cipher and scrambling encryption, and the secure retrieval of images in the YUV color space, respectively. These foundational methods set the stage for the development of more complex retrieval mechanisms, emphasizing the need for encrypted images to be both secure and retrievable.

Building on these encryption methodologies, studies [25–29] delve into advanced retrieval mechanisms that operate within the constraints of encrypted do-mains. Huffman-code based retrieval, secure Local Binary Pattern (LBP) features, and the application of Markov processes exemplify the innovative approaches taken to extract meaningful information from encrypted images. These mechanisms showcase the progression towards retrieval systems that are not only secure but also capable of accurately identifying images similar to a query image.

Application-specific solutions, as discussed in works [30–33], further demonstrate the usefulness and practical implications of privacy-preserving image retrieval technologies. The introduction of privacy-preserving Scale-Invariant Feature Transform (SIFT), Convolutional Neural Network (CNN) frame-works for medical data, and systems like PIC (Privacy-preserving Image search system on Cloud) high-light the field's move towards addressing specific needs such as medical diagnosis aid and large-scale image search with fine-grained access control. These tailored approaches underline the importance of developing encrypted image retrieval systems that accommodate to the unique requirements of different domains, emphasizing the critical role of privacy in sensitive applications.

Lastly, comparative analyses and the proposal of new frameworks, as seen in [34–42], reflect the ongoing evolution and refinement of privacy-preserving image retrieval methods. By comparing homomorphic encryption-based techniques with feature/index randomization-based techniques and introducing novel frameworks like IES-CBIR for outsourced storage and retrieval, these works contribute to a deeper under-standing of the trade-offs between security, efficiency, and usability. The development of dynamic verifiable retrieval schemes and multi-indexed hashing approaches represents the cutting edge of research, aimed at improving the precision and security of image retrieval in cloud-based systems. We present a comparative analysis between each of the mentioned methods in Table I. It is found that non-of the existing methods have used VGG16 for encrypted CBIR. Furthermore, to the best of our knowledge, none of the existing works on Content-Based Image Retrieval (CBIR) have optimized the trained model through post-training quantization. Quantization techniques are pivotal for augmenting the computational efficiency of CBIR systems. By diminishing the bit-precision of neural network parameters, quantization paves the way for swifter and more energy-efficient computations across extensive hardware accelerators. This precision reduction not only accelerates computational speed but

also aids in energy conservation, rendering it an indispensable strategy for the deployment of large-scale CBIR systems.

As presented in Fig. 1, the general taxonomy of existing works in the field of Content-Based Image Retrieval (CBIR) can be categorized into four main areas, each focusing on different aspects and method-ologies to enhance privacy-preserving image retrieval. The first category, Encryption Methods for Image Retrieval, aims to develop and apply encryption techniques to secure image data while maintaining its retrievability. Approaches in this category include symmetric and asymmetric encryption schemes, homomorphic encryption for computation on encrypted images, and watermarking techniques for securely embedding retrieval information.

TABLE I. SUMMARY OF METHODS AND FEATURES IN ENCRYPTED CONTENT-BASED IMAGE RETRIEVAL (CBIR) AND RELATED APPLICATIONS

Work	Method of Encryption	Type of Features Extracted	Retrieval Model	Quantization
[24]	Permutation of DCT coefficients	Histogram at each frequency position	Unsupervised/Supervised retrieval using integrated distances and conditional probabilities	×
[25]	Stream cipher and scrambling encryption of DC and AC coefficients	AC coefficients histogram	Statistical comparison of histograms to return closest encrypted images	×
[27]	Stream cipher, permutation cipher for Huffman code	Encrypted Huffman-code histograms	Feature comparison to return similar content images	×
[29]	Color value substitution, block permutation, intra-block pixel permutation	Normalized histogram of encrypted visual words (BOEW model)	Direct similarity measurement between feature vectors on the cloud server side	×
[28]	Big-block permutation, polyalphabetic cipher	Secure Local Binary Pattern (LBP) features	Retrieval based on order-preserving encryption and secure LBP features	×
[43]	Stream cipher and permutation encryption for DCT coefficients	Transition probability matrices (Markov process), SVM classification	Feature extraction and classification to evaluate similarity between encrypted images	×
[30]	Homomorphic encryption	Privacy-preserving SIFT features	Secure SIFT feature extraction and representation in the encrypted domain	×
[32]	Homomorphic encryption	Wavelet-based image features	Secure CBIR for diagnosis aid systems with data confidentiality preservation	×
[33]	Not explicitly mentioned (uses encrypted images)	Features extracted by CNN	Privacy-preserving classification and retrieval using CNN framework	×
[40]	Homomorphic encryption vs. feature/index randomization	Not explicitly mentioned	Comparative analysis for confidentiality-preserving image search	×
[34]	Randomized binary encoding and Gaussian random matrix	Encrypted visual words	Secure index construction for large-scale image retrieval without decryption	×
[42]	Comparable encryption	Not explicitly mentioned	Encrypted image search scheme balancing efficiency and privacy	×
[37]	Feature descriptors extracted by CNN models, encrypted hierarchical index tree	Features extracted by CNN, hierarchical index tree	Similarity search for encrypted images using SEI with enhanced key privacy	×
[38]	Pre-trained CNN model, encrypted index based on K-means clustering	Features extracted by CNN	Dynamic verifiable retrieval over encrypted images (DVREI) scheme	×
[36]	ViT model and ITQ method for multi-indexed hashing (MIH)	Features extracted by ViT model, secure Hamming distance protocol	Privacy-preserving content-based image retrieval using MIH	×
[41]	Cryptographic techniques and secure indexing schemes	Not explicitly mentioned	Content-based retrieval over encrypted multimedia databases	×
[35]	Based on IES-CBIR scheme	Not explicitly mentioned	Outsourced privacy-preserving storage and retrieval in large shared image repositories	×
[44]	Encryption of difference matrices of RGB components (EDH-CBIR)	Euclidean distance between feature vectors to measure similarity	Euclidean distance	×
[45]	Permutation of big-blocks and substitution of binary code of DCT coefficients	Local Markov Features and Bag-of-Words Model	Efficient Content-Based Image Retrieval (CBIR) service for image owners with direct feature extraction	×
[46]	Thumbnail-preserving encryption (TPE)	HSV and uniform local binary pattern (ULBP) features	Thumbnail-Preserving Encryption (TPE)	×
[47]	Fully Homomorphic Encryption (FHE), Secure Multi-Party Computation (SMPC), AES, FPE	Local invariant features	CBIR, Attribute-Based Encryption (ABE), Functional Encryption (FE)	×
[48]	Asymmetric scalar-product-preserving encryption (ASPE)	Local invariant features	Invariant Features Selection	×
[49]	Watermark-based encryption	Dominant local patterns, Relative directional edge patterns (RDEP)	Dominant Local Patterns, Watermark Encryption	×
Ours	Multi-Tiered Texture and Color Encryption (MTTCE)	VGG-16 based features	Efficiency aware VGG-16 retrieval model with supporting three quantization modes	√

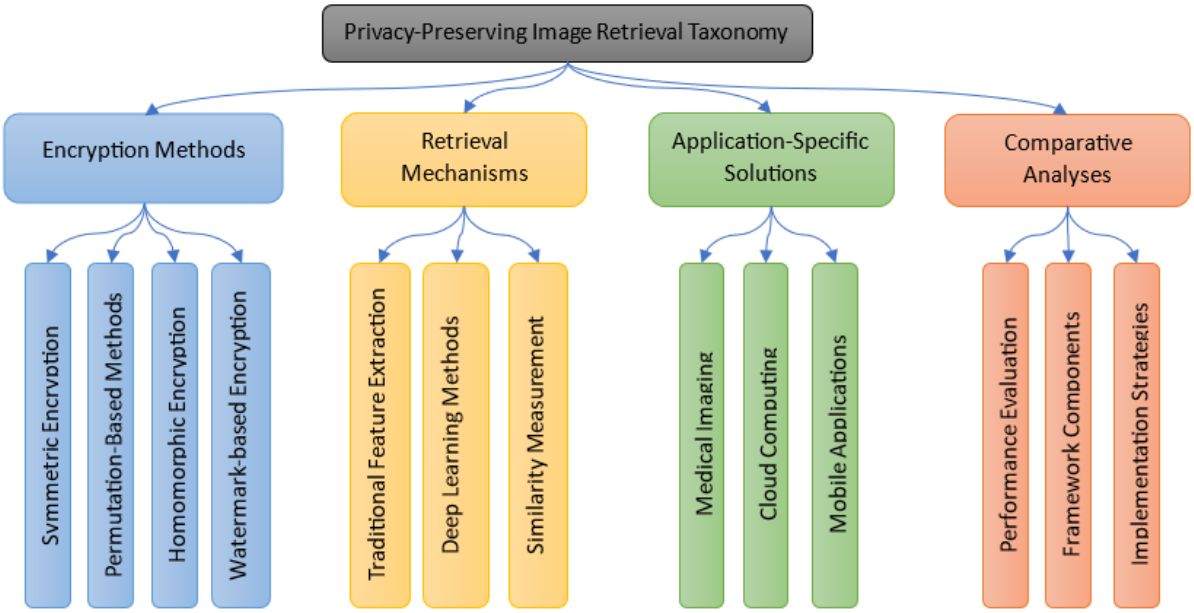

Fig. 1. The general taxonomy of existing works focuses on CBIR.

The second category, Retrieval Mechanisms and Feature Extraction, seeks to enhance the efficiency and accuracy of retrieval mechanisms by optimizing feature extraction and similarity measures. This involves traditional feature extraction methods such as SIFT, SURF, and HOG, as well as deep learning-based feature extraction using convolutional neural networks (CNNs) and similarity measures including Euclidean distance, cosine similarity, and advanced metrics tailored for encrypted data.

The third category focuses on Application-Specific Solutions for Privacy-Preserving Retrieval, developing tailored solutions that address specific application domains and their unique privacy requirements. The fourth and final category, Comparative Analyses and Framework Proposals, involves conducting comparative studies of existing methods and proposing comprehensive frameworks to guide future research and implementation in the field of CBIR. This taxonomy provides a structured overview of the diverse approaches and innovations aimed at enhancing the privacy and effectiveness of image retrieval systems.

B. Efficiency Optimization Methods

The quest for optimizing search efficiency in large-scale image databases has given rise to several innovative strategies in the realm of content-based image retrieval (CBIR). Among these, quantized deep learning frameworks have become a cornerstone for hashing-based retrieval systems. Hashing Nets tackles the challenge of limited storage space and computational resources, especially relevant for satellite remote sensing image retrieval and unmanned aerial vehicles (UAVs). The proposed Quantized Deep Learning to Hash (QDLH) framework introduces binarized weights and activation functions, creating a lightweight neural network that considerably reduces the demand on hardware resources [50]. This need for efficiency extends to other works, such as the application of deep model quantization and compression to CNNs on ASIC chips, achieving comparable performance to floating-point models with a mere 2-bit weights quantization [51].

Addressing the constraints of mobile device capabilities, the OMCBIR framework emerges with an innovative solution. It presents ALNet, an ultra-lightweight neural network that harnesses pointwise group convolution, channel shuffle, and a convolutional attention module, substantially minimizing the model's size without compromising on retrieval accuracy [52].

For the complex domain of remote sensing images, a different approach is introduced in (2021) with the APQ method. This method leverages a multi-scale attention-based CNN combined with an enhanced product quantization technique to efficiently compress features, significantly improving retrieval performance [53].

Furthermore, the domain of multi-label image retrieval presents its own set of challenges, particularly in maintaining semantic integrity within quantization errors. The Multi-label Contrastive Hashing (MCH) technique innovates with a curriculum strategy that carefully adjusts the quantization loss weight, fostering the preservation of multi-level semantic similarity more effectively than prior hashing-based retrieval methods [54].

C. Research Gap

Despite the remarkable progress in content-based image retrieval (CBIR) systems, there exists a visible gap in the domain of medical image retrieval, specifically concerning the post-training optimization of pre-trained models. Current literature prominently features methods optimized for general image databases or remote sensing images, focusing on quantization and hashing to enhance retrieval efficiency and reduce computational demands. These advancements, while substantial, are not directly translatable to the unique requirements and complexity of medical imaging.

Medical images, such as radiographs, MRI, and CT scans, present distinct challenges due to their high-dimensional data and the critical need for precision in feature extraction to capture clinically relevant information. Pre-trained models, often developed on natural image datasets, which may not naturally capture the specific features and diseases that are noticeable in medical images. Fine-tuning pre-trained models on medical datasets is a common strategy to transfer learning from one domain to another; however, there is an obvious absence of research into post-training optimization techniques that refine these models further for the medical domain.

Existing fine-tuning practices largely focus on adapting the pre-trained models to new datasets by retraining some layers while keeping others frozen. This method is beneficial but does not fully exploit the potential of the models to conform to the particularities of medical image analysis. Post-training optimization can involve techniques such as neural architecture search (NAS) tailored to medical datasets, advanced quantization specifically sensitive to medical imaging features, or specialized regularization strategies that address the overfitting risks associated with medical image datasets, which are often smaller and more variable than those used in the training of general models.

The absence of a dedicated post-training optimization phase means that while the models may perform well on general benchmarks, their efficacy in medical scenarios, where the margin for error is minimal, could be significantly enhanced. Such optimizations could lead to improvements in retrieval accuracy, relevance of retrieved images, and ultimately, clinical usefulness. Our article aims to investigate this gap, proposing a framework for post-training optimization of pre-trained models specifically fine-tuned for medical image retrieval, with the goal of maximizing the clinical relevance and accuracy of the retrieved results.

IV. METHODOLOGY

A. Problem Formulation

Given an encrypted image database and a feature extraction model, the system must efficiently process query images to retrieve and return images that are visually similar to the query image. This process involves encryption and decryption of images and features to ensure privacy preservation, requiring the system to address several key challenges:

1) Efficient and secure encryption and decryption mechanisms that maintain the usability and accessibility of the retrieval system.

2) The development of a robust feature extraction model that can effectively operate on encrypted images to extract meaningful features for identification.

3) The implementation of an effective identification and retrieval mechanism that can operate in the encrypted domain, ensuring privacy while maintaining high retrieval accuracy.

4) The development of a quantization of the trained retrieval model to enable efficiency and scalability.

5) The main objective is to achieve a balance between privacy preservation, computational efficiency, and retrieval effectiveness within a cloud-based CBIR system.

More formally, we write the problem based on the following entities and roles:

1) Data owner: The data owner possesses an image database $I = \{i_1, i_2, \dots, i_n\}$ consisting of n images. Each image in the database is encrypted using a unique key from the set $K = \{k_1, k_2, \dots, k_n\}$, resulting in an encrypted image database $E = \{e_1, e_2, \dots, e_n\}$. The data owner also trains a feature extraction model Ψ with E, which is capable of extracting features from encrypted images. The encrypted database E and the model Ψ are then uploaded to the cloud server for storage and deployment.

2) Cloud server: Serving as the backbone for storage and computational power, the cloud server stores the encrypted image database E and the feature extraction model Ψ. Upon receiving an encrypted query image EQ from a query user, the server processes E and EQ through Ψ to extract their features, $F_E = \{f_{e1}, f_{e2}, \dots, f_{en}\}$for the database images and F_{EQ} for the query image. The similarity between F_E and F_{EQ} is assessed using the Euclidean distance metric to identify the k most similar images. The identifiers of these images, $ER = \{er_{ID1}, er_{ID2}, \dots, er_{IDk}\}$, are then returned to the query user.

3) Query user: The query user is interested in retrieving images similar to a query image Q. The query image is first encrypted to EQ and then uploaded to the cloud server. After receiving the encrypted query result set ER from the cloud server, the query user sends these identifiers $ID^R = \{ID_{R1}, ID_{R2}, \dots, ID_{Rk}\}$ back to the data owner to obtain the corresponding decryption keys $RK = \{rk_{IDR1}, rk_{IDR2}, \dots, rk_{IDRk}\}$.

These keys allow the query user to decrypt the received images, resulting in the final retrieval set $R = \{r_{IDR1}, r_{IDR2}, \dots, r_{IDRk}\}$. We present the mathematical symbols used in this article in Table II.

TABLE II. THE SUMMARY OF NOTATIONS

Notations	Definitions
n	The size of the image dataset
m^2	Number of blocks
$I = \{i_1, i_2, \cdots, i_n\}$	The plaintext image dataset
$E = \{e_1, e_2, \cdots, e_n\}$	The encrypted image dataset
$K = \{k_1, k_2, \dots, k_n\}$	The set of security keys
$F_E = \{f_{e1}, f_{e2}, \cdots, f_{en}\}$	The encrypted image feature dataset
Q	The plaintext query image
EQ	The encrypted query image
F_{EQ}	The encrypted query image feature
Ψ	The feature extraction model
$\mathrm{ER} = \{\mathrm{er}_{\mathrm{ID1}}, \mathrm{er}_{\mathrm{ID2}}, \cdots, \mathrm{er}_{\mathrm{IDk}}\}$	The encrypted query result images
$\mathrm{ID}_R = \{\mathrm{ID}_{R1}, \mathrm{ID}_{R2}, \cdots, \mathrm{ID}_{Rk}\}$	The query result image ID
$\mathrm{RK} = \{\mathrm{rk}_{\mathrm{IDR\ 1}}, \mathrm{rk}_{\mathrm{IDR\ 2}}, \cdots, \mathrm{rk}_{\mathrm{IDRk}}\}$	The key corresponding to the resulting image
$R = \{r_{\mathrm{IDR1}}, r_{\mathrm{IDR2}}, \cdots, r_{\mathrm{IDRk}}\}$	The query result image
$\mathrm{sub}\,I = \{\mathrm{sub}\,I_1, \mathrm{sub}\,I_2, \cdots, \mathrm{sub}\,I_{m^2}\}$	Image subblock
$\mathrm{sub}\,E = \{\mathrm{sub}\,E_1, \mathrm{sub}\,E_2, \cdots, \mathrm{sub}\,E_{m^2}\}$	Encrypted image subblock

B. System Architecture

The sequence diagram showed in Fig. 2 presents the operational workflow of a privacy-preserving Content-Based Image Retrieval (CBIR) system, illustrating the interactions among three key entities: the Data Owner, the Cloud Server, and the Query User. The process begins with the Data Owner, who possesses a database of images intended for secure retrieval. This entity takes the initial steps by encrypting the image database, creating an encrypted image database (E), and developing a feature extraction model (Ψ). These components are crucial for ensuring privacy and facilitating feature-based image retrieval in an encrypted domain. Once prepared, the Data Owner uploads both the encrypted image database and the feature extraction model to the Cloud Server, a platform that provides the necessary storage and computational power for the system. Upon the system's readiness to handle queries, the Query User engages by first encrypting a query image (EQ) using a similar encryption methodology as the Data Owner. This encrypted query image is then uploaded to the Cloud Server, indicating the start of the retrieval process. The Cloud Server, leveraging the previously uploaded feature extraction model, processes both encrypted query image and encrypted image database to extract their respective features. Although the specific approach for measuring similarity is abstracted in this diagram, the Cloud Server identifies the k most similar images to the query and returns their identifiers (ER) to the Query User. The interaction between the Query User and the Data Owner is reinitiated when the Query User requests the decryption keys for the received images. The Data Owner responds by providing the necessary decryption keys (RK), enabling the Query User to decrypt and access the final retrieval set of images (R). This sequence diagram effectively encapsulates the secure and private workflow of a CBIR system, emphasizing encryption for privacy, cloud-based feature extraction and retrieval, and decryption for accessing the retrieved images, all while abstracting the complexities of similarity measurement and feature extraction details.

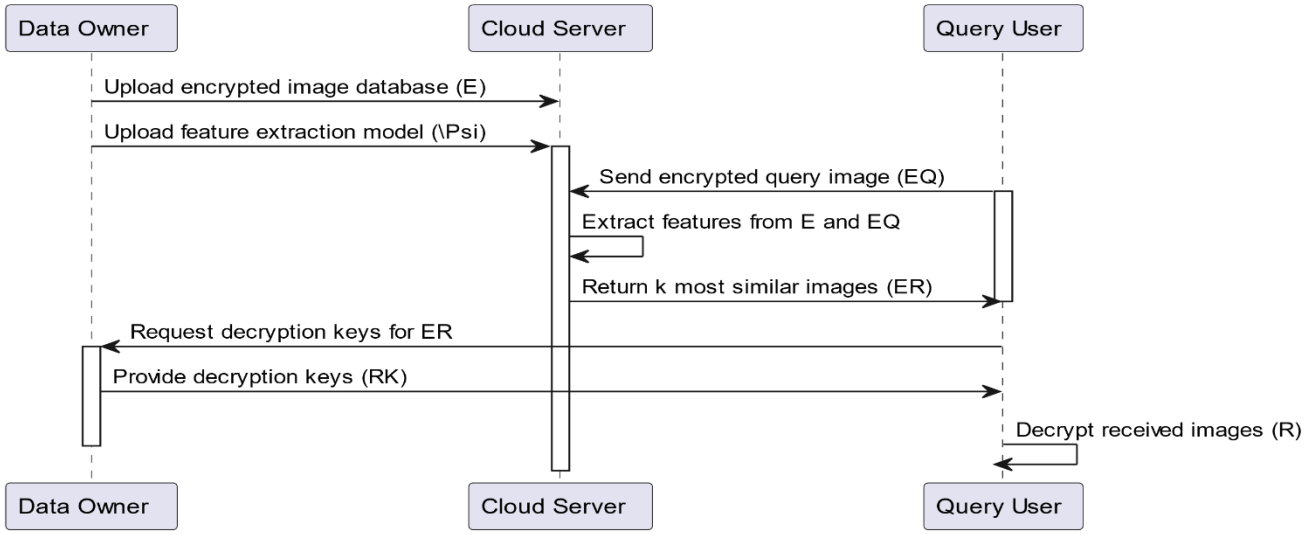

Fig. 2. Sequence diagram illustrating the operational flow of a privacy-preserving Content-Based Image Retrieval (CBIR) system.

C. Framework

The framework proposed in Fig. 3 provides a holistic view of encrypted image retrieval with modified VGG-16 architecture and state-of-the-art quantization methods. It helps in solving the dual issues of ensuring privacy through encryption and enabling efficient content-based image retrieval.

The process started with the encryption of data, all images from the training distribution were then encrypted and stored in cloud. This step makes certain that data privacy is met at every point of retrieval. The basis of the framework is a VGG-16 feature extraction module adjusted to work with encrypted images. Using the encrypted input image, convolutional layers and max pooling is applied on top of this leading to fully connected layers that give an output feature vector.

The same framework also introduces a key model optimization module just to keep the model's memory and efficiency. There are four quantization methods evaluated: Max Quantization, KL Quantization, 99 Percentile Quantization method and Full Model Method. These techniques were applied to the original model weights to obtain a quantized model that trades retrieval accuracy for computational and memory efficiency.

The quantized model is followed by a fine-tuning step on the encrypted dataset preparing its parameters, then to suit the encrypted medical images. The tuned model is then deployed to cloud where given query images while returns the K similar images in database.

The framework also covers secure user interaction flow as well. The users are able to send encrypted query images, receive the encrypted results and ask for decryption keys from data owners to obtain the recognized multimedia files. This means the secure conversion of text throughout its lifecycle as it is fetched.

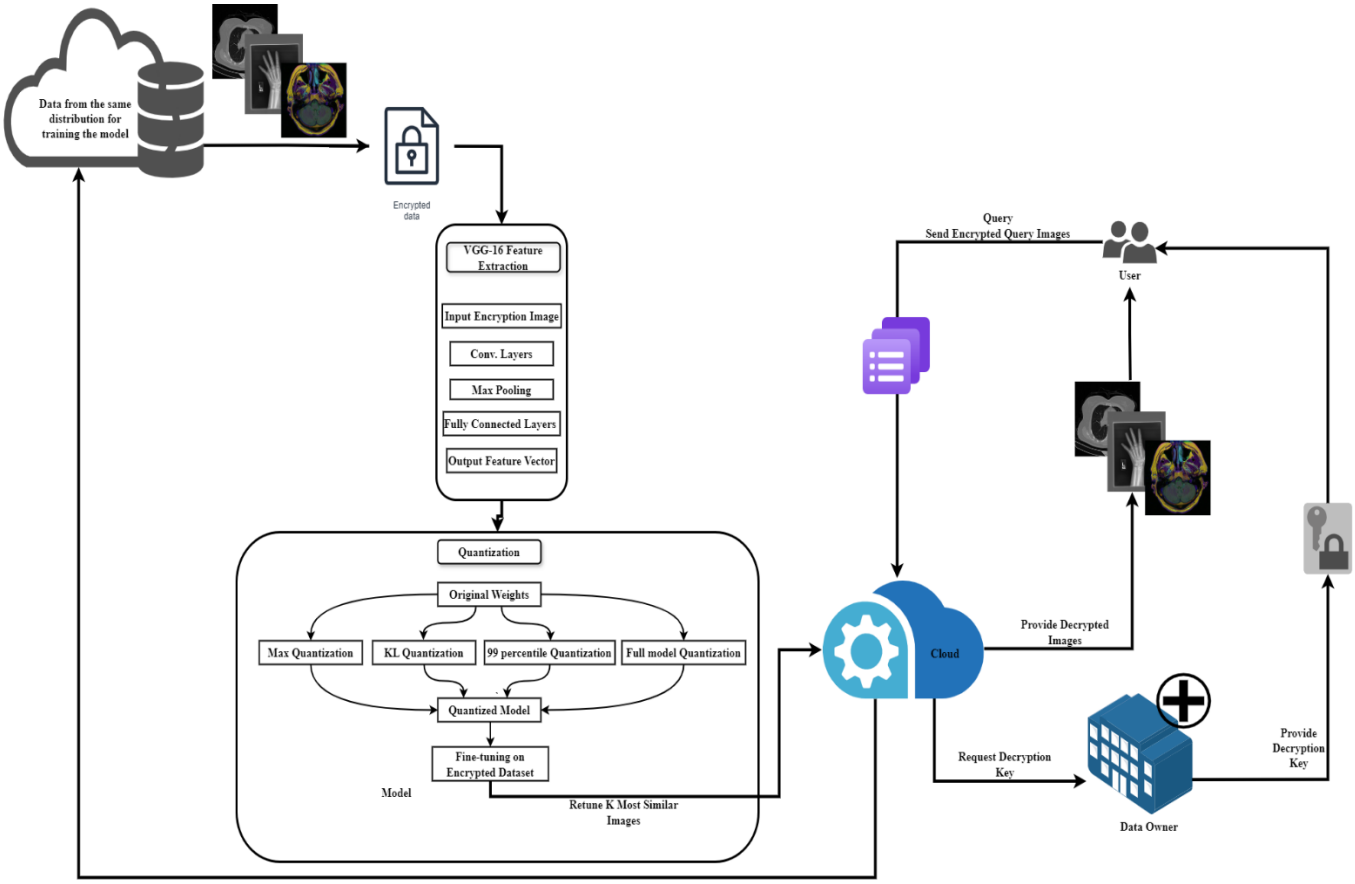

Fig. 3. Encrypted medical image retrieval framework using VGG-16 and quantization techniques.

In addition to support for advanced encryption, feature extraction and quantization as well secure user interactions this framework presents one of the most state-of-the-art tools in privacy-preserving content-based image retrieval especially useful e.g., when it comes to dealing with sensitive domains like medical imaging.

D. Encryption

The encryption algorithm detailed in the provided pseudocode is a comprehensive multi-stage process designed to enhance the security of content-based image retrieval (CBIR) systems. It ensures both local and global image information is encrypted, safeguarding essential retrieval features while preventing unauthorized access. By methodically encrypting local textures, global textures, and color information, this algorithm offers a robust solution to maintaining the privacy and security of image data in CBIR applications. The provided pseudocode outlines a multi-stage encryption algorithm designed to enhance the security of content-based image retrieval (CBIR) systems. The primary objective is to secure both local and global information of an image, ensuring that features necessary for retrieval are preserved while the image remains protected from unauthorized access. The algorithm begins with the input of an original image, denoted as I. The output of the algorithm is the encrypted image I_{enc} and the encryption key K, which is essential for decrypting the image. The start of the process involves dividing the original image I into non-overlapping subblocks, referred to as B_i . This segmentation is crucial as it sets the stage for detailed encryption targeting local textures within the image. In the first stage of encryption, the algorithm focuses on obscuring local texture information within each subblock. For every subblock B_i, the positions of the RGB channel values are scrambled. These scrambling hides local textures, making it difficult for unauthorized users to recognize patterns that could lead to information leakage. Next, the algorithm shifts to protect global texture information. This is accomplished by randomly scrambling the positions between the subblocks B_i. By shuffling these blocks, the spatial relationships within the image are altered, further confusing any attempts to understand the encrypted image's structure. The final stage of encryption deals with securing global color information. For each scrambled subblock B_i, the algorithm substitutes the RGB channel values and swaps the channels. This step not only secures the color information but also ensures that the value substitution is fixedly related to the position of the encrypted subblocks, adding an additional layer of complexity to the encryption scheme. To complete the encryption process, the algorithm generates an encryption key K using a built-in random function. This key is essential for the decryption procedure, allowing authorized users to reverse the encryption process and recover the original image. The processed subblocks are then combined to form the encrypted image $I_{\text{enc.}}$. The encrypted image is now secured, containing no recognizable original content, and can only be decrypted with the correct encryption key. The end of the algorithm marks the completion of the encryption process.

Algorithm 1: Multi-Stage Image Encryption for CBIR Systems

1. **Input:** Original Image I
2. **Output:** Encrypted Image I_{enc}, Encryption Key K
3. Start
4. Divide the original image I into non-overlapping subblocks B_i
5. **for** each subblock B_i **do**
6. Scramble the positions of the RGB channel values within B_i to hide local texture information
7. **end for**
8. Randomly scramble the positions between subblocks B_i to obscure global texture information
9. **for** each scrambled subblock B_i **do**
10. Substitute the RGB channel values and swap the channels to secure global color information
11. **end for**
12. Generate the encryption key K using a built-in random function
13. Combine the processed subblocks to form the encrypted image I_{enc}
14. **End**

It would summarize the encryption algorithm steps well and in a logical, flowing manner, the multistage operation will tend to be secure to CBIR systems more as it will protect the local histogram features and unlink global texture color information. Intrablock scrambling, interblock scrambling and channel substitution are combined to give strong encryption that can be used to provide protection from access to the image data without a key yet enable integrity for the features of useful image retrieval.

As an illustrating example, we present the results of applying the methods to medical images, showcasing a hand X-ray, an MRI brain scan, and a skull X-ray as depicted in Fig. 4. The original images show distinct medical details essential for diagnosis: the hand X-ray reveals the bone structure, the MRI scan distinguishes soft tissue contrasts within the brain, and the skull X-ray clearly outlines the facial bone anatomy. In their encrypted state, these images transform into a mosaic of indistinguishable colored blocks, effectively covering any diagnostic information and ensuring data privacy. The decrypted images, when compared with the originals, show no apparent loss of detail or quality, indicating that the encryption process is fully reversible and maintains the integrity of medical information. This demonstrates the encryption method's potential for securing sensitive health information as well as efficient storage and distribution of medical images are feasible while maintaining their reversibility for medical use. The encryption approach, therefore, strikes a balance between protecting patient confidentiality and preserving the utility of medical images for diagnostic purposes. This highlights the robustness of the encryption method in terms of its practical use in healthcare settings, where patient data protection is essential without limiting healthcare professionals' capability to examine and evaluate medical imaging as needed.

The security analysis of the algorithm is presented in the following steps:

Step 1: Dividing the Image into Subblocks

Let I be the original image of size $M \times N$. The image is divided into non-overlapping subblocks B_i of size $m \times n$, where $m \times n$ is a divisor of $M \times N$. Let k be the number of subblocks, so

$$k = \frac{M \times N}{m \times n} \quad (1)$$

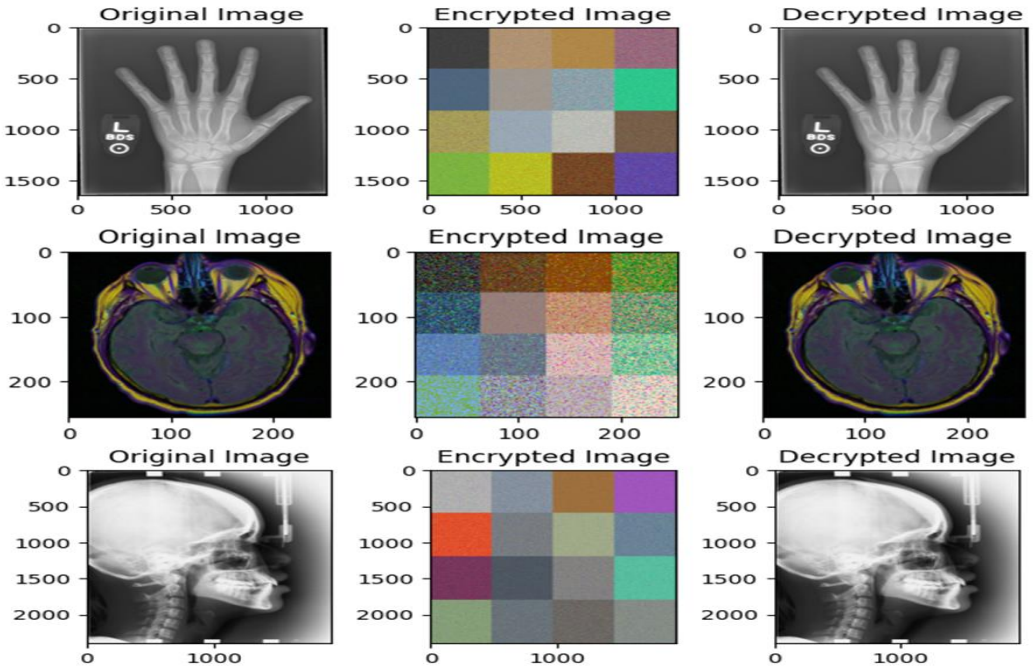

Fig. 4. Encryption process illustrated: original, encrypted, and decrypted states of medical images for secure image retrieval.

Step 2: Scrambling the RGB Values Within Each Subblock

For each subblock B_i, the positions of the RGB channel values are scrambled. The number of possible permutations for each subblock is $(m \times n)$!. Given k subblocks, the total number of possible permutations is:

$$(m \times n)!^k \quad (2)$$

This large permutation space ensures that local textures are effectively obscured.

Step 3: Scrambling the Positions of Subblocks

After local texture encryption, the subblocks are shuffled. The number of possible permutations of k subblocks is !. Combining this with the permutations from the previous step, the total number of possible configurations is:

$$(m \times n)!^k \times k! \quad (3)$$

This significantly increases the complexity, making it difficult for unauthorized users to reconstruct the original image based on subblock positions.

Step 4: Substituting RGB Values and Swapping Channels

For each subblock, the RGB channel values are substituted, and the channels are swapped. Suppose each RGB value can be replaced with any other value within the range $[0,255]$.

The number of possible substitutions for each channel is 256. Since there are three channels, the total number of possible substitutions for each subblock is:$(256)^3$

Swapping the channels adds an additional 3 ! permutations.

Therefore, the total number of configurations for each subblock considering color information is: Combining all the steps, the total number of possible permutations for the entire image encryption process is:

$$(m \times n)!^k \times k! \times (256)^3 \times 3! \quad (4)$$

This enormous number represents the total permutation space, which is computationally infeasible to brute force.

The encryption key K is generated using a built-in random function. The key must be sufficiently long and complex to cover the permutation space generated by the above steps. Assume the key length is L. The entropy of the key is:

$$(K) = L\log_2(N) \quad (5)$$

where, N is the number of possible values for each part of the key. A sufficiently large L ensures that the key space is large enough to resist brute-force attacks.

The security of this encryption algorithm is primarily based on the vast permutation space created by scrambling subblock positions, substituting RGB values, and swapping channels. The combined permutations create an extremely large key space, making brute-force attacks impractical. The algorithm's design ensures that both local and global features of the image are secured, providing robust protection against unauthorized access while maintaining the integrity necessary for effective CBIR system retrieval.

E. VGG-16

The VGGNet architecture, a highly influential convolutional neural network (CNN), was developed through a partnership between the Visual Geometry Group at the University of Oxford and Google DeepMind. It represents a significant evolution in the CNN landscape, building upon the foundational principles established by its predecessor, AlexNet. VGGNet has become well-known for its architectural depth and the use of uniformly small convolutional filters, specifically 3x3 kernels, which have set a new standard for feature extraction in image recognition tasks. As shown in Fig. 5, the architecture consists of a repeating pattern of convolution layers followed by 2x2 max pooling layers, maintaining the spatial hierarchy of the features being learned. This approach not only maintains the spatial hierarchy of the features being learned but also allows for an increase in the depth of the network without a corresponding explosion in computational complexity. By deepening the network, VGGNet significantly enhances the hierarchical feature learning process, capturing fine-grained details that are often crucial for accurate image classification. The VGGNet family comprises several models, among which the VGGNet-16 and VGGNet-19 are the most notable. These models are differentiated by the number of weight layers they contain: 16 and 19, respectively. Both have demonstrated remarkable performance on large-scale image recognition tasks, contributing to their widespread adoption in the field. The depth of these networks has proven to be a key factor in their ability to perform complex image classifications with high accuracy, making them particularly useful in applications where precision is critical. Furthermore, the impact of VGGNet extends beyond its immediate performance. The architecture has provided invaluable insights into the design of deep neural networks, influencing subsequent innovations in the domain. Its widespread use as a pre-trained model for a variety of tasks points out its significance because it provides a starting point for additional optimizing and adaptation for specialized applications, such as those in medical imaging and other fields where detailed feature detection is crucial [55–57].

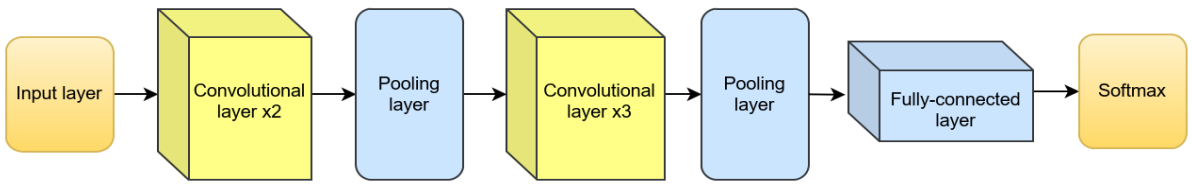

Fig. 5. Architectural flowchart of a deep convolutional neural network with repeating convolution and pooling layers, culminating in a softmax output.

F. Quantization

Content-Based Image Retrieval (CBIR) is a technique for retrieving relevant images from a database based on a given query image. When dealing with encrypted images, this task becomes more complex. To address this, we propose a method using a VGG-16 model with various quantization techniques to improve efficiency and performance. The first step involves applying quantization to the last layer of a pre-trained VGG-16 model. Quantization reduces the precision of the weights, leading to a smaller model size and faster computations. The quantization methods used include max quantization, KL quantization, 99 percentile quantization, and full model quantization. For max quantization, the maximum absolute value of the weights is calculated, and each weight is divided by this value. In KL quantization, the histogram of the weights is computed, and quantization thresholds are determined using Kullback-Leibler divergence, followed by quantizing the weights based on these thresholds. For 99 percentile quantization, the 99th percentile of the absolute weights is calculated, and weights are divided by this value. Full quantization applies quantization to all layers of the model. After quantization, the original weights of the last layer are replaced with the quantized weights, resulting in a quantized model. The quantization pseudocode is presented in Algorithm 2.

Algorithm 2: Quantization for Last Layer

1: **Input**: Trained model M, quantization mode *mode*
2: **Output**: Quantized model M'
3: W← weights of the last layer of M
4: **if** *mode* == "max" **then**
5: $W_{max} \leftarrow \max(|W|)$
6: Quantize W as $W_q = W / W_{max}$
7: **else if** *mode* == "KL" **then**
8: Compute the histogram of W
9: Determine the quantization thresholds using KL divergence
10: Quantize W using the determined thresholds
11: **else if** *mode* == "99%" **then**
12: $W_{99} \leftarrow 99^{th}$ percentile of $|W|$
13: Quantize W as $W_q = W / W_{99}$
14: **else if** *mode* == "fully" **then**
15: Quantize all layers of M
16: **end if**
17: Replace the last layer weights of M with W_q
18: **return** M'

Algorithm 3: Encrypted CBIR with VGG-16 Quantized Models

1: **Input:** Encrypted image dataset D, query image Q, quantization mode *mode*
2: **Output:** Retrieved images
3: Initialization:
4: Load pre-trained VGG-16 model M
5: Apply Quantization for Last Layer algorithm with mode\text{mode}mode
6: Fine-tune the quantized model Mon encrypted images
7: Perform image retrieval using the quantized model M'
8: **return** Retrieved images

G. Big O Notation

Quantization is a technique used to reduce the precision of the weights in a neural network, which in turn reduces the memory footprint and computational requirements of the model. Here's a complexity analysis of the memory usage after applying different quantization methods to the last layer of the VGG-16 model as described in the algorithms.

Let's denote the number of weights in the last layer of the VGG-16 model as W. Typically, these weights are stored as 32-bit floating-point numbers (i.e., each weight takes 4 bytes).

The memory usage for the last layer before quantization is:

$$\text{Μεμορψ}_{\beta\varepsilon\phi o\rho\varepsilon} = W \times 4\beta\psi\tau\varepsilon\sigma \quad (6)$$

Quantization reduces the precision of these weights, typically to 8-bit integers (i.e., each weight takes 1 byte). The memory usage after quantization depends on the number of weights and the quantization method used.

Quantization reduces the precision of these weights, typically to 8-bit integers (i.e., each weight takes 1 byte). The memory usage after quantization depends on the number of weights and the quantization method used.

Max quantization scales the weights based on the maximum absolute value. The quantized weights are stored as 8-bit integers, and an additional scaling factor needs to be stored (usually as a 32-bit float).

The memory usage for max quantization is:

$\text{Memory}_{\max} = W \times 1$ byte +4 bytes (scaling factor)

KL quantization uses histograms and thresholds determined using KL divergence. Similar to max quantization, the weights are stored as 8-bit integers, and additional parameters for thresholds may be stored (assuming negligible storage for histograms and thresholds compared to the number of weights).

H. KL Quantization

KL quantization uses histograms and thresholds determined using KL divergence. Similar to max quantization, the weights are stored as 8-bit integers, and additional parameters for thresholds may be stored (assuming negligible storage for histograms and thresholds compared to the number of weights).

The memory usage for KL quantization is:

$\text{Memory}_{\text{KL}} = W \times 1$ byte +4 bytes (scaling factor)

99 Percentile Quantization

99 percentile quantization scales the weights based on the 99th percentile of their absolute values. Again, weights are stored as 8-bit integers with an additional scaling factor.

The memory usage for 99 percentile quantization is:

$\text{Memory}_{99\%} = W \times 1$ byte +4 bytes (scaling factor)

Full quantization applies to all layers of the model. For simplicity, let's denote the total number of weights in the VGG-16 model as $W_{\text{total.}}$.

The memory usage for full model quantization is:

$$\text{Memory}_{\text{full}} = W_{\text{total}} \times 1 \text{ byte} + \text{scaling factors}$$

If there are L layers in the VGG-16 model, and each layer has a scaling factor, the memory usage for scaling factors is negligible compared to the total number of weights.

The memory usage complexity remains linear in terms of the number of weights, but the actual memory usage is significantly reduced due to the lower precision (8-bit vs. 32-bit). In summary, quantization effectively reduces the memory usage of the model by a factor of 4 (since 32-bit weights are converted to 8-bit weights), while the complexity in terms of the number of weights remains the same. This reduction is crucial for deploying models in resource-constrained environments and for speeding up computations during inference. In summary, quantization effectively reduces the memory usage of the model by a factor of 4 (since 32-bit weights are converted to 8-bit weights), while the complexity in terms of the number of weights remains the same. This reduction is crucial for deploying models in resource-constrained environments and for speeding up computations during inference.

V. Experimental Results and Analysis

The fundamental components and their parametric relationships in the proposed framework are systematically presented in Table III. The framework's architecture comprises three main parameter categories: encryption parameters, VGG-16 model configurations, and quantization methods. The encryption parameters, particularly the block size, represent a crucial design choice that determines the balance between security strength and feature preservation. Similarly, the RGB channel substitution parameter provides flexibility in controlling the degree of visual information protection while maintaining essential image characteristics for retrieval purposes. In the deep learning component, the VGG-16 model parameters - learning rate and batch size - establish the foundation for stable model training and resource utilization. The quantization parameters offer different approaches to model compression, each with its unique characteristics: max quantization prioritizes compression efficiency, 99% percentile quantization aims for balanced preservation of significant features, and KL divergence quantization focuses on maintaining statistical distributions. This systematic organization of parameters provides system designers with a clear framework for making informed decisions based on their specific requirements for privacy preservation, computational efficiency, and retrieval capability.

The study leveraged three distinct datasets, each encompassing 1200 medical images. The datasets were further segregated for training and validation, and testing purposes, employing a 1000:100:100 split, respectively. Notably, the datasets were categorized based on anatomical regions, specifically the lung, brain, and bone. This stratified and anatomically categorized data structure facilitated the robust evaluation of the proposed method across diverse medical imaging domains, we present our experimental design in Table IV.

TABLE III. Experimental Design for Comparing our Proposed Method with Benchmark

Parameter Category	Parameter	Value/Range	Trade-offs	Recommended Setting
Encryption	Block Size	1-8 blocks	- Larger blocks: Better retrieval, lower security - Smaller blocks: Higher security, poor retrieval	8 blocks for medical applications requiring balance of privacy and accuracy
	RGB Channel Substitution	0-255	Security vs. Feature quality	Application-dependent, moderate values (128-192) for balanced performance
VGG-16	Learning Rate	0.001	Training speed vs. Stability	0.001 (demonstrated optimal convergence)
	Batch Size	32	Memory usage vs. Training stability	32 (balances resource usage and stability)
Quantization	Max Quantization	Maximum value in layer	Compression vs. Accuracy	When maximum compression is needed
	99% Percentile	99th percentile value	Moderate compression with optimal accuracy	Default choice for most applications
	KL Divergence	Distribution-based	Computational complexity vs. Precision	When distribution preservation is critical

TABLE IV. Experimental Design for Comparing our Proposed Method with Benchmark

	VIT	VGG16
Experiments	Plaintext- encrypted image where block size (1,8)	Plaintext- encrypted image where block size (1, 8)
Pretrained	TRUE	TRUE
Num Classes	3	3
Optimizer	Adam	Adam
Learning Rate	0.001	0.001
Loss Function	CrossEntropyLoss	CrossEntropyLoss
Epochs	100	100
Batch Size	32	32

A. Encryption Quality Analysis

For the experimental evaluation of the encryption, we compare the encryption algorithm adopted with the widely recognized AES standard as shown in Table V. The comparison is grounded on a set of metrics that typically measure the efficacy of encryption methods in terms of security and the ability to resist statistical analysis.

- Entropy: Our algorithm records an entropy of 7.392, which indicates a substantial amount of randomness, although slightly less than the optimal value of 8. This suggests that while the algorithm introduces randomness, there may be room for improvement. The AES, in an unusual turn, shows an entropy value extremely close to zero, which typically would suggest a lack of randomness; however, this could imply a perfect encryption where the output is indistinguishable from a completely random source.
- MSE (Mean Squared Error): In this evaluation, a higher MSE between the original and encrypted images indicates stronger encryption. The MSE for our algorithm is significantly higher than that of AES, suggesting that our method may provide a more robust alteration of pixel values, thereby potentially increasing security.
- SSIM (Structural Similarity Index): The SSIM for our algorithm is noticeably higher than that of AES, implying that the structural integrity of the image is somewhat retained. In contrast, AES's lower SSIM reinforces its role as a robust encryption standard by substantially altering the image structure to secure the data effectively.
- PSNR (Peak Signal-to-Noise Ratio): The lower PSNR associated with our algorithm complements the high MSE, affirming the extensive alteration from the original image. Meanwhile, AES exhibits a higher PSNR, which is usually indicative of a decryption process that maintains image quality. However, in the encryption phase, a lower PSNR may be more desirable as it indicates a greater level of distortion.

B. Image Retrieval Performance Analysis

The confusion matrices and the classification metrics presented in Fig. 6(a), 6(b), and 6(c) illustrates the performance of a Vision Transformer (ViT) based image retrieval system under different conditions: plaintext images and images encrypted using 8 blocks and 1 block encryption. These matrices provide insight into how encryption affects the model's ability to correctly classify images into one of three categories: bone, chest, and MRI.

TABLE V. COMPARATIVE EVALUATION OF ENCRYPTION METRICS: OUR ALGORITHM VERSUS AES

Metric	Our Algorithm	AES
Entropy	**7.392129887**	1.4426951601859516e-10
MSE	**12766.02**	2625.48
SSIM	0.273204	**0.00283838**
PSNR	**7.070247655**	13.93870288

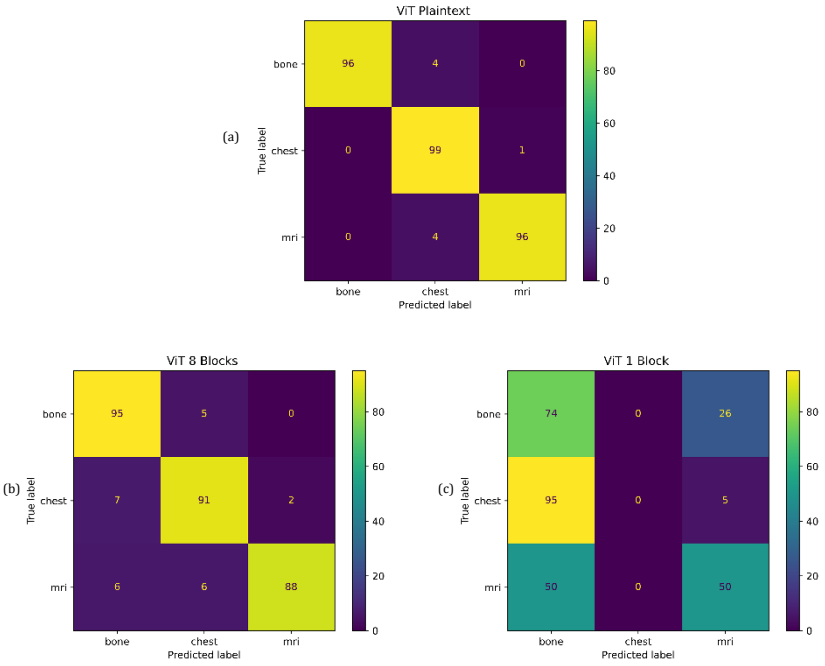

Fig. 6. Confusion matrix of ViT based image retrieval (a) plaintext (b) encrypted domain using 8 block encryption (c) encrypted domain using 1 block encryption.

In the plaintext scenario, the performance of the ViT model is well-exemplified since correctness can be identified with great precision for each category in three classes. It is particularly very high for the classes of bone and MRI, of which only few images can misclassify. The chest category provides a high accuracy impression too but slightly messes up with the remaining categories. A high true positive rate for each class actually represents that the model was well tuned for the features of each class.

The 8 blocks encryption has somewhat degraded performance. It still maintains relatively high accuracy, though an increase in misclassification is seen, especially with the bone and chest categories. A minimal decrease in accuracy is observed in the MRI category. This hints that most probably the model starts having difficulty extracting features when receiving high-encryption images. That is one possible reason for the weakness in performance, considering that encryption puts noise in images, thus blurring features that our model depends on for classification.

With 1 block encryption, the performance drops significantly, most notably in the bone and MRI categories. The model's ability to correctly identify bone images is drastically reduced, and there is a considerable increase in the misclassification of MRI images as bone or chest. The SSIM for this encryption level is likely quite low, indicating that the structural integrity of the images is heavily compromised, and the model is unable to extract meaningful features for accurate classification.

The gradual performance decline across the matrices from plaintext to 1 block encryption illustrates the trade-off between security and usability in encrypted domains. It underscores the challenge of maintaining feature extraction capabilities for classification tasks while also securing the images against unauthorized access or analysis. This balance is crucial in applications where both the confidentiality of the data and the accuracy of automated systems are of paramount importance.

1) ViT-based image retrieval: Table VI presents overall classification metrics for a Vision Transformer (ViT) based image retrieval system, comparing performance across three different categories (Bone, Chest, MRI) and three models: ViT Plaintext, ViT 8 Blocks, and ViT 1 Block.

ViT Plaintext exhibits exceptional performance with accuracy rates hovering around 98% for all categories. The precision is perfect for 'Bone' and nearly so for 'MRI', indicating almost no false positives for these categories. 'Chest' has a slightly lower precision, which suggests a few more false positives but still maintains a high recall, indicating it successfully identified most true 'Chest' cases.

ViT 8 Blocks shows a drop in performance across all metrics, which is expected as the encryption level increases. Accuracy for 'Bone' and 'Chest' categories decreases by about 4-5%, and for 'MRI' by 3%, compared to the plaintext model. Precision sees a more notable decline, particularly for 'Bone' where it drops by over 12%. This suggests that the 8-block encryption introduces enough ambiguity to affect the model's ability to correctly identify features specific to 'Bone' images.

'MRI' retains high precision but suffers in recall, indicating that while most 'MRI' predictions are correct, the model fails to identify all 'MRI' images, likely due to feature loss in the encryption process. ViT 1 Block shows a significant decrease in performance.

For 'Bone', the accuracy and precision are notably lower, and the recall is moderately high, which may indicate a higher number of false negatives. In 'Chest', precision is undefined, which occurs when the denominator in the precision calculation is zero; this happens if there were no predictions made for the 'Chest' category or all predictions were incorrect. The recall for 'Chest' is 0%, confirming the model did not correctly identify any 'Chest' images. For 'MRI', both precision and recall have decreased significantly, with accuracy being marginally better than 'Bone' but still substantially lower than in the other models.

2) VGG-16-based image retrieval: As provided in Fig. 7(a), 7(b) and 7(c), The confusion matrices and classification results for the VGG-16-based image retrieval present a comprehensive view of the model's performance across three different settings: plaintext, 8-block encryption, and 1-block encryption domains.

TABLE VI. OVERALL CLASSIFICATION METRICS OF VIT BASED IMAGE RETRIEVAL

Model	Category	Accuracy (%)	Precision (%)	Recall (%)
ViT Plaintext	Bone	98.67	100.00	96.00
	Chest	97.00	92.52	99.00
	MRI	98.33	98.97	96.00
ViT 8 Blocks	Bone	94.00	87.96	95.00
	Chest	93.33	89.22	91.00
	MRI	95.33	97.78	88.00
ViT 1 Block	Bone	43.00	33.79	74.00
	Chest	66.67	Undefined	0.00
	MRI	73.00	61.73	50.00

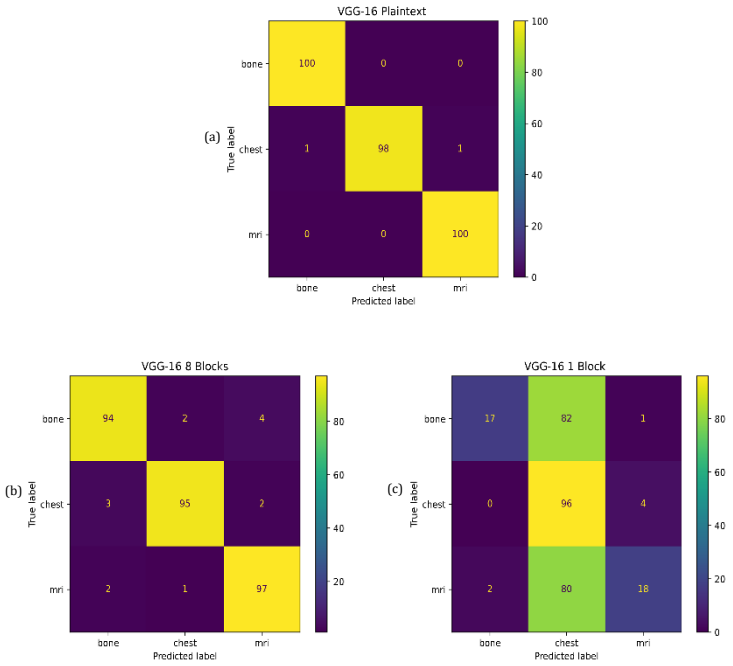

Fig. 7. Confusion matrix of VGG-16 based image retrieval (a) plaintext (b) encrypted domain using 8 block encryption (c) encrypted domain using 1 block encryption.

In the plaintext scenario, the VGG-16 model demonstrates near-perfect classification accuracy, with an impressive 99.33% accuracy across all categories. The precision and recall rates are equally outstanding for the 'Bone' and 'MRI' categories, both hitting 100% in recall, indicating that every relevant image was correctly retrieved. 'Chest' images also show a high level of precision and recall, indicating that the model can distinguish between these medical images with high reliability in an unencrypted domain.

Moving to the 8-block encryption domain, there is a slight but noticeable decrease in performance. The accuracy remains high at 95.33%, yet there are marginal drops in precision and recall for the 'Bone' category, indicating a slight increase in both false positives and false negatives. The Chest and MRI classes lost little accuracy, indicating that encryption adds some ambiguity, however, the model can identify features for retrieval quite effectively.

The 1-block encryption domain delivers quite different findings. The accuracy dropped to 43.67%, indicating that at such high levels of encryption, the model was unable to classify images correctly. Interestingly, the 'Chest' category shows a high recall, suggesting that while the model can identify 'Chest' images, it does so with a high rate of false positives as reflected in the lower precision rate. 'Bone' and 'MRI' categories exhibit poor recall rates, indicating a majority of relevant images are missed, yet when they are identified, they tend to be correct, as shown by the higher precision rates.

Overall, the matrices and results in Table VII highlight the challenges posed by encryption on the ability of CBIR systems to maintain high retrieval performance. They underscore the need for specialized approaches to manage encrypted image data, especially as the strength of encryption increases and significantly impacts the extraction of features critical for accurate image retrieval.

TABLE VII. OVERALL CLASSIFICATION METRICS OF VGG-16 BASED IMAGE RETRIEVAL

Model	Category	Accuracy (%)	Precision (%)	Recall (%)
VGG-16 Plaintext	Bone	99.33	99.01	100.00
	Chest	99.33	100.00	98.00
	MRI	99.33	99.01	100.00
VGG-16 8 Blocks	Bone	95.33	94.95	94.00
	Chest	95.33	96.94	95.00
	MRI	95.33	94.17	97.00
VGG-16 1 Block	Bone	43.67	89.47	17.00
	Chest	43.67	37.21	96.00
	MRI	43.67	78.26	18.00

3) Quantized VGG-16 based image retrieval: In evaluating the performance of the quantized VGG-16 models for image retrieval, we can interpret the provided confusion matrices for each quantization technique as shown in Fig. 8(a), 8(b), 8(c) and 8(d):

a) VGG-16 8 Blocks Final Layer Quantized (Max): The model shows high performance with most 'Bone', 'Chest', and 'MRI' images correctly identified, evidenced by the high numbers on the diagonal of the confusion matrix. There is a slight confusion between the 'Bone' and 'MRI' categories, and to a lesser extent with 'Chest'. However, with 93 out of 100 correct predictions for both 'Bone' and 'Chest' categories, the model demonstrates robustness under the max quantization method.

b) VGG-16 8 Blocks Final Layer Quantized (KL): The model performance is slightly reduced under KL quantization, particularly in the 'Chest' category, where we see a decrease to 91 correct predictions and an increase in misclassification as 'MRI'. The 'MRI' category also shows a minor decrease in performance, with 95 correct predictions. This suggests that the KL quantization may introduce some information loss that affects the model's classification ability.

c) VGG-16 8 Blocks Final Layer Quantized (99%): Under the 99% percentile quantization, the model's performance improves, with 'Bone' and 'MRI' classifications seeing an uptick in correct predictions to 95, and 'Chest' holding steady at 93. This indicates that the 99% percentile method effectively maintains crucial information for image classification.

d) VGG-16 8 Blocks Fully Quantized: The fully quantized model reveals a drastic change in the confusion pattern, with most 'Bone' and 'Chest' images being misclassified as 'MRI'. While this may appear to be a major decline in model performance, it is worth noting that such a high degree of quantization is likely to significantly reduce the model's size and computational requirements. This trade-off between size and accuracy may be beneficial in specific applications where computational efficiency is prioritized over classification accuracy.

Table VIII presents classification metrics for a VGG-16 model with the final layer quantized under different schemes: maximum, KL divergence, 99%, and fully quantized, across medical image categories (Bone, Chest, MRI). The performance metrics detailed—accuracy, precision, and recall—provide insights into the impact of these quantization methods on the model's ability to effectively process encrypted medical images segmented into 8 blocks.

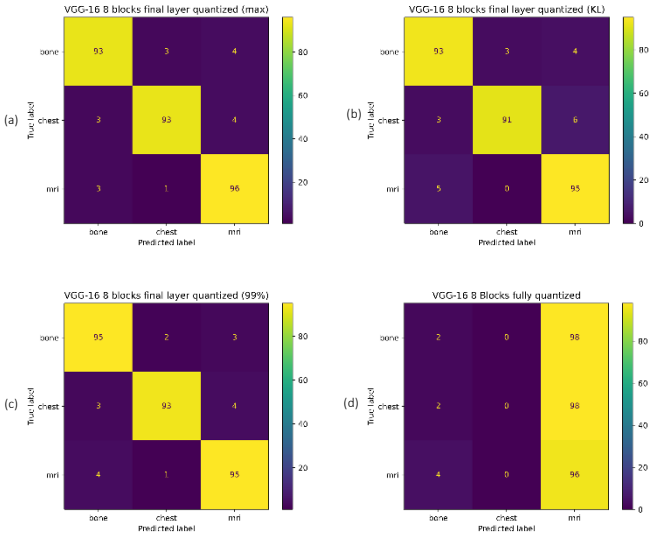

Fig. 8. Confusion matrix of Quantized VGG-16 based image retrieval (a) max (b) KL encryption (c) 99% percentile (d) fully quantized.

TABLE VIII. CLASSIFICATION METRICS OF QUANTIZED VGG-16 FOR DIFFERENT FINAL LAYER QUANTIZATION TYPES MAXIMUM, KL, 99%, AND FULLY QUANTIZED

Model	Category	Accuracy (%)	Precision (%)	Recall (%)
VGG-16 8 blocks final layer quantized (max)	Bone	0.94	0.939	0.93
	Chest	0.94	0.958	0.93
	MRI	0.94	0.923	0.96
VGG-16 8 blocks final layer quantized (KL)	Bone	0.93	0.920	0.93
	Chest	0.93	0.968	0.91
	MRI	0.93	0.904	0.95
VGG-16 8 blocks final layer quantized (99%)	Bone	0.946	0.931	0.959
	Chest	0.946	0.968	0.93
	MRI	0.946	0.940	0.95

For models where only the final layer is quantized using the maximum, KL divergence, and 99% methods, the results indicate relatively robust performance with minor variations across the different types of quantization. Accuracy remains consistently high, around 0.94 to 0.946, under the maximum and 99% quantization types, with slight decreases in the KL quantized version to 0.93. This minimal variance suggests that moderate quantization of the final layer does not drastically affect the model's ability to classify medical images accurately.

Precision and recall metrics show more variation. In the maximum quantization scenario, precision is high across all categories but slightly lower for MRI at 0.923, while recall is highest for MRI at 0.96, indicating good sensitivity. Under KL quantization, precision peaks for Chest images at 0.968, whereas recall is slightly reduced for Chest and MRI. The 99% quantization appears to balance both precision and recall effectively, particularly showing a notable improvement in Bone recall to 0.959.

However, a stark contrast is observed when the model is fully quantized. The accuracy decreases to 0.326 in all categories, while precision and recall also see a substantial decline, except for MRI recall which surprisingly stays high at 0.96. A significant decrease in performance highlights the challenges and possible drawbacks of full model quantization, especially in complex applications like as medical image classification, where accuracy and responsiveness to features are important.

The analysis underscores the necessity of careful consideration in the extent of quantization applied within neural networks, especially in medical settings where accuracy and reliability are crucial. Moderate quantization of the final layer preserves functionality, while full quantization severely impairs performance, suggesting that a balanced approach is essential for maintaining the efficacy of encrypted image retrieval systems.

VI. Discussion

In this discussion, we delve into the intricacies of implementing the VGG-16 model for encrypted and quantized image retrieval, focusing on the trade-offs between security, efficiency, and effectiveness. We explore the optimization of encryption and quantization methods to achieve optimal performance in a Content-Based Image Retrieval (CBIR) system.

1) The VGG-16 model was selected for the image retrieval task due to its deep architecture, which is highly effective in capturing complex textural details and spatial hierarchies in images. This characteristic is especially beneficial in medical imaging, where precise feature extraction is crucial for accurate classification and diagnosis. VGG-16's use of small, 3x3 convolutional kernels allows for a deeper network that can learn a rich hierarchy of features at multiple scales, making it a robust choice for both plaintext and encrypted image retrieval systems. The model's well-documented success in various image recognition tasks supports its adoption in exploring the impacts of encryption and quantization on medical image analysis.

2) The Role of Encryption and the Effectiveness and Security Balance Encryption plays a critical role in ensuring the confidentiality and security of sensitive data, such as medical images, in a CBIR system. However, as demonstrated by the performance metrics across different encryption granularities, there is a significant trade-off between security and the effectiveness of image retrieval. While high encryption levels (e.g., 1-block encryption) offer stronger data protection, they severely degrade the quality of features available for the model to learn and identify, leading to poorer classification accuracy and recall rates. Conversely, lighter encryption (e.g., 8-blocks) provides less security but maintains higher usability of the data, allowing for more effective feature extraction and image retrieval. Balancing these aspects requires a strategic approach to encryption that considers the specific needs for security versus the operational requirements for image analysis and retrieval.

3) Optimization of Encryption for Achieving Required Performance To optimize encryption for performance, it is essential to tailor the encryption methods to the specific characteristics of the data and the requirements of the retrieval system. For medical images, where detail and accuracy are paramount, adopting an encryption approach that preserves more structural and textural integrity—such as region-specific encryption or varying encryption levels based on the sensitivity of the image content—might be more suitable. Techniques such as homomorphic encryption could also be explored, as they allow certain types of computations to be carried out on encrypted data, thus maintaining privacy without sacrificing the ability to perform effective image retrieval. Additionally, optimizing the trade-off between block size in encryption and feature extraction needs is crucial, as smaller blocks increase security but reduce the quality of features, impacting model performance.

4) The Role of Quantization and the Efficiency and Effectiveness Trade-Off Quantization addresses the need to reduce the computational complexity and storage requirements of deep learning models, making them more suitable for deployment in resource-constrained environments such as mobile devices or in cloud-based architectures where computational efficiency is critical. The quantization of VGG-16, particularly at the final layer, has demonstrated the possibility of maintaining relatively high accuracy and precision with minor trade-offs in performance. However, fully quantized models exhibit a significant drop in performance, highlighting the delicate balance between efficiency and effectiveness. Quantization introduces noise and approximation errors into the model, which can degrade the accuracy if not managed carefully. The choice of quantization technique—maximum, KL divergence, or percentile-based—impacts how much of the critical information is preserved and thus the overall effectiveness of the model. Striking a balance between reducing computational demands and maintaining high classification accuracy is essential, particularly when dealing with high-stakes applications like medical diagnostics.

VII. Conclusion and Future Work

The fundamental problem of this study is in the development of a Content-Based Image Retrieval system that can effectively optimize the trade-off between privacy for encrypted images and computational efficiency and retrieval effectiveness. This is of utmost significance, particularly in fields where privacy and accuracy are closely related, such as in medical imaging. The proposed system is designed to efficiently handle an encrypted database of images, prioritizing the extraction of features and retrieval of images.

The primary objective of this study has been to investigate the VGG-16 architecture, which has gained recognition as a very efficient deep convolutional network for extracting high-level features from few imagine samples. Two essential modifications for preparing the model for encrypted domains were making adjustments to the granularity and applying quantization techniques in the last layer of the model. The primary objective of these approaches is to investigate the feasible and controlled impacts of encryption and quantization on the performance of a model.

Research findings indicated that VGG-16, in general, maintained a high level of accuracy in classifying plaintext images, but its performance declined as the level of detail arose. Therefore, it is the encryption with a granularity of 1-block that significantly and severely affects the ability of this model in accurately classifying images into particular classifications such as bone or MRI. The impact of quantization was considerably less severe but it became quite significant when the entire model was completely quantized. Research findings indicate that although a certain level of encryption and quantization can possibly be achieved with no impact on performance, over use of these techniques results in significant performance impairment.

One of the primary limitations of this study is the focus on a single model architecture (VGG-16) and a limited set of encryption and quantization strategies. Future research could explore other architectures like ResNet or EfficientNet, which might interact differently with encryption and quantization. Additionally, experimenting with newer encryption techniques like homomorphic encryption could offer insights into less disruptive methods. Further exploration into adaptive encryption and quantization strategies that dynamically adjust based on the content sensitivity and retrieval needs could also enhance system performance and usability.

Moreover, incorporating more sophisticated metrics for assessing the quality of retrieved images, such as structural similarity indexes or perceptual quality measures, could provide a deeper understanding of how encryption and quantization affect the perceived quality of images. Future studies should also consider the implementation of these systems in real-world applications, evaluating their practicality and efficiency in operational environments.

Acknowledgment

Communication of this research is made possible through monetary assistance by Universiti Tun Hussein Onn Malaysia and the UTHM Publisher's Office via Publication Fund E15216.

References

[1] S. Sikandar, R. Mahum, and A. Alsalman, A novel hybrid approach for a content-based image retrieval using feature fusion, vol. 13, no. 7, p. 4581, https://doi.org/10.3390/app13074581.

[2] G. Sumbul, J. Kang, and B. Demir, Deep learning for image search and retrieval in large remote sensing archives, pp. 150–160, https://doi.org/10.1002/9781119646181.ch11.

[3] M. S. Sayed, A. A. A. Gad-Elrab, K. A. Fathy, and K. R. Raslan, Unsupervised Content Based Image Retrieval Using Pre-Trained CNN and PCNN Features Extractors., vol. 16, no. 1, doi: 10.22266/ijies2023.0228.50.

[4] M. A. Mohammed, M. A. Hussain, Z. A. Oraibi, Z. A. Abduljabbar, and V. O. Nyangaresi, Secure Content Based Image Retrieval System Using Deep Learning, https://doi.org/10.56714/bj-rs.49.2.9.

[5] Z. Xia, L. Lu, T. Qiu, H. J. Shim, X. Chen, and B. Jeon, A Privacy-Preserving Image Retrieval Based on AC-Coefficients and Color Histograms in Cloud Environment., vol. 58, no. 1, doi:10.32604 /cmc.2019.02688.

[6] S S. H. Wady and R. Z. Yousif, A Secure Medical Image Transmission System Based On 2D Logistic Map and Diffie-Hellman Key Exchange Mechanisms, vol. 6, no. 2, pp. 94–104, https://doi.org/10.21928/uhdjst.v6n2y2022.pp94-104.

[7] S. R. Nair, High and Low-Level Classification based Features for Content-based Image Retrieval in Privacy-Preserving, https://doi.org/10.212-03/rs.3.rs-1461397/v1.

[8] F. Zhou, S. Qin, R. Hou, and Z. Zhang, Privacy-preserving image retrieval in a distributed environment, vol. 37, no. 10, pp. 7478–7501, https://doi.org/10.1002/int.22890.

[9] C. Zhang, C. Xu, K. Sharif, and L. Zhu, Privacy-preserving contact tracing in 5G-integrated and blockchain-based medical applications, vol. 77, p. 103520, https://doi.org/10.1016/j.csi.2021.1035-20.

[10] T. Janani and M. Brindha, Privacy-preserving transfer learning-based secure quantum image retrieval in encrypted domain for cloud environment, vol. 32, no. 2, p. 23003, https://doi.org/10.1117/1.JEI.32.2.023003.

[11] M. Naveen, A Review on Content Based Image Retrieval System Features Derived by Deep Learning Models, http://dx.doi.org/10.22214/ijraset.2021.39172.

[12] W. Wu, D. Li, J. Du, X. Gao, W. Gu, F. Zhao, X. Feng and H. Yan, An Intelligent Diagnosis Method of Brain MRI Tumor Segmentation Using Deep Convolutional Neural Network and SVM Algorithm, vol. 2020, p. 6789306, https://doi.org/10.1155/2020/6789306.

[13] Y. Zhang, Y. Ma, and Y. Liu, Convolution-Bidirectional Temporal Convolutional Network for Protein Secondary Structure Prediction, vol. 10, pp. 117469–117476, doi: 10.1109/ACCESS.2022.3219490.

[14] Y. Zhao and C. Feng, Remote Sensing Image Water Body Recognition Algorithm Based on Deep Convolution Generating Network and Combined Features, vol. 2022, p. 9932251, doi: 10.1155/2022/9932251.

[15] E. M. Alsaedi and A. kadhim Farhan, Retrieving Encrypted Images Using Convolution Neural Network and Fully Homomorphic Encryption, doi: 10.21123/bsj.2022.6550.

[16] Y. Wang, F. Wang, F. Liu, and X. Wang, Securing content-based image retrieval on the cloud using generative models, vol. 81, no. 22, pp. 31219–31243, https://doi.org/10.1007/s11042-022-12880-6.

[17] P. Desai, J. Pujari, C. Sujatha, A. Kamble, and A. Kambli, Hybrid approach for content-based image retrieval using VGG16 layered architecture and SVM: an application of deep learning, vol. 2, no. 3, p. 170, https://doi.org/10.1007/s42979-021-00529-4.

[18] C. I. Kerley, Y. Huo, S. Chaganti, S. Bao, M. B. Patel, and B. A. Landman, Montage Based 3D Medical Image Retrieval From Traumatic Brain Injury Cohort Using Deep Convolutional Neural Network, doi: 10.1117/12.2512559.

[19] A. Ahmed, A. O. Almagrabi, and A. H. Osman, Pre-Trained Convolution Neural Networks Models for Content-Based Medical Image Retrieval, doi: 10.21833/ijaas.2022.12.002.

[20] D. B. Mahesh, B. Madhuri, and R. L. D, Integration of Optimized Local Directional Weber Pattern With Faster Region Convolutional Neural

Network for Enhanced Medical Image Retrieval and Classification, doi: 10.1111/coin.12506.

[21] C. Palai, P. K. Jena, S. R. Pattanaik, T. Panigrahi, and T. Mishra, Content-Based Image Retrieval Using Encoder Based RGB and Texture Feature Fusion, doi: 10.14569/ijacsa.2023.0140328.

[22] C. G. Sotomayor, M. Mendoza, V. Castañeda, H. Farías, G. T. Molina, G. Pereira, S. Härtel, M. Solar and M. Araya, Content-Based Medical Image Retrieval and Intelligent Interactive Visual Browser for Medical Education, Research and Care, doi: 10.3390/diagnostics11081470.

[23] S. Bilquees, H. Dawood, H. Dawood, N. Majeed, A. Javed, and M. T. Mahmood, Noise Resilient Local Gradient Orientation for Content-Based Image Retrieval, vol. 2021, p. 4151482, doi: 10.1155/2021/4151482.

[24] X. Zhang and H. Cheng, Histogram-based retrieval for encrypted JPEG images, pp. 446–449, doi: 10.1109/ChinaSIP.2014.6889282.

[25] H. Cheng, X. Zhang, and J. Yu, AC-coefficient histogram-based retrieval for encrypted JPEG images, vol. 75, no. 21, pp. 13791–13803, doi: 10.1007/s11042-015-2741-z

[26] P. Li and Z. Situ, Encrypted JPEG image retrieval using histograms of transformed coefficients, no. November, pp. 1140–1144, doi: 10.1109/APSIPAASC47483.2019.9023179.

[27] H. Liang, X. Zhang, and H. Cheng, Huffman-code based retrieval for encrypted JPEG images, vol. 61, pp. 149–156, doi: 10.1016/j.jvcir.2019.03.021.

[28] Z. Xia, L. Wang, J. Tang, N. N. Xiong, and J. Weng, A Privacy-Preserving Image Retrieval Scheme Using Secure Local Binary Pattern in Cloud Computing, vol. 8, no. 1, pp. 318–330, doi: 10.1109/TNSE.2020.3038218.

[29] Z. Xia, L. Jiang, D. Liu, L. Lu, and B. Jeon, BOEW: A Content-Based Image Retrieval Scheme Using Bag-of-Encrypted-Words in Cloud Computing, vol. 15, no. 1, pp. 202–214, doi: 10.1109/TSC.2019.2927215

[30] C.-Y. Hsu, C.-S. Lu, and S.-C. Pei, Image feature extraction in encrypted domain with privacy-preserving SIFT, vol. 21, no. 11, pp. 4593–4607, doi: 10.1109/TIP.2012.2204272.

[31] L. Zhang, T. Jung, P. Feng, K. Liu, X. Y. Li, and Y. Liu, PIC: Enable large-scale privacy preserving content-based image search on cloud, vol. 2015-Decem, pp. 949–958, doi: 10.1109/ICPP.2015.104.

[32] R. Bellafqira, G. Coatrieux, D. Bouslimi, and G. Quellec, Content-based image retrieval in homomorphic encryption domain, in 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015, pp. 2944–2947, doi: 10.1109/EMBC.2015.7319009.

[33] C. Guo, J. Jia, K. K. R. Choo, and Y. Jie, Privacy-preserving image search (PPIS): Secure classification and searching using convolutional neural network over large-scale encrypted medical images, vol. 99, p. 102021, doi: 10.1016/j.cose.2020.102021.

[34] B. Cheng, L. Zhuo, Y. Bai, Y. Peng, and J. Zhang, Secure index construction for privacy-preserving large-scale image retrieval, no. 4142009, pp. 116–120, doi: 10.1109/BDCloud.2014.36.

[35] B. Ferreira, J. Rodrigues, J. Leitão, and H. Domingos, Practical Privacy-Preserving Content-Based Retrieval in Cloud Image Repositories, vol. 7, no. 3, pp. 784–798, doi: 10.1109/TCC.2017.2669999.

[36] J. Huang, Y. Luo, M. Xu, S. Fu, and K. Huang, Accelerating Privacy-Preserving Image Retrieval with Multi-Index Hashing, pp. 492–497, doi: 10.1109/SEC54971.2022.00075.

[37] Y. Li, J. Ma, Y. Miao, Y. Wang, X. Liu, and K. K. R. Choo, Similarity Search for Encrypted Images in Secure Cloud Computing, vol. 10, no. 2, pp. 1142–1155, doi: 10.1109/TCC.2020.2989923.

[38] Y. Li, J. Ma, Y. Miao, H. Li, Q. Yan, Y. Wang, X. Liu and K. K. R. Choo, DVREI: Dynamic Verifiable Retrieval Over Encrypted Images, vol. 71, no. 8, pp. 1755–1769, doi: 10.1109/TC.2021.3106482.

[39] D. Liu, J. Shen, Z. Xia, and X. Sun, A content-based image retrieval scheme using an encrypted difference histogram in cloud computing, vol. 8, no. 3, p. 96, https://doi.org/10.3390/info8030096.

[40] W. Lu, A. L. Varna, and M. Wu, Confidentiality-preserving image search: A comparative study between homomorphic encryption and distance-preserving randomization, vol. 2, pp. 125–141, doi: 10.1109/ACCESS.2014.2307057.

[41] W. Lu, A. Swaminathan, A. L. Varna, and M. Wu, Enabling search over encrypted multimedia databases, in Media Forensics and Security, International Society for Optics and Photonics, 2009, p. 725418, doi : 10.1117/12.806980.

[42] Q. Zou, J. Wang, and X. Chen, Secure Encrypted Image Search in Mobile Cloud Computing, pp. 572–575, https://doi.org/10.1109/BWCCA.201-5.41.

[43] H. Cheng, X. Zhang, J. Yu, and F. Li, Markov process-based retrieval for encrypted JPEG images, vol. 2016, no. 1, pp. 1–9, doi: 10.1186/s13635-015-0028-6.

[44] D. Liu, J. Shen, Z. Xia, and X. Sun, A content-based image retrieval scheme using an encrypted difference histogram in cloud computing, vol. 8, no. 3, p. 96, https://doi.org/10.3390/info8030096.

[45] P. Yu, J. Tang, Z. Xia, Z. Li, and J. Weng, A Privacy-Preserving JPEG Image Retrieval Scheme Using the Local Markov Feature and Bag-of-Words Model in Cloud Computing, vol. 11, no. 3, pp. 2885–2896, doi: 10.1109/TCC.2022.3233421.

[46] Y. Ma, X. Chai, Z. Gan, and Y. Zhang, Privacy-preserving TPE-based JPEG image retrieval in cloud-assisted internet of things, doi: 10.1109/JIOT.2023.3301042.

[47] K. Jayashree, Secureimagesec: A privacy-preserving framework for outsourced picture representation with content-based image retrieval, no. Preprint, pp. 1–22, doi: 10.3233/IDA-240265.

[48] S. Kumar, A. K. Pal, S. K. H. Islam, and M. Hammoudeh, Secure and efficient image retrieval through invariant features selection in insecure cloud environments, vol. 35, no. 7, pp. 4855–4880, https://doi.org/10.1007/s00521-021-06054-y.

[49] G. Sucharitha, D. Godavarthi, J. V. N. Ramesh, and M. I. Khan, Secure and efficient content-based image retrieval using dominant local patterns and watermark encryption in cloud computing, pp. 1–17, https://doi.org/10.1007/s10586-024-04635-9.

[50] P. Li, L. Han, X. Tao, X. Zhang, C. Grecos, A. Plaza and P. Ren, Hashing Nets for Hashing: A Quantized Deep Learning to Hash Framework for Remote Sensing Image Retrieval, vol. 58, no. 10, pp. 7331–7345, doi: 10.1109/TGRS.2020.2981997.

[51] B. Yang, L. Yang, X. Li, W. Zhang, H. Zhou, Y. Zhang, Y. Ren and Y. Shi, 2-bit Model Compression of Deep Convolutional Neural Network on ASIC Engine for Image Retrieval, https://doi.org/10.48550/arXiv.1905.03362.

[52] X. Zhang, C. Bai, and K. Kpalma, OMCBIR: Offline mobile content-based image retrieval with lightweight CNN optimization, vol. 76, p. 102355, doi: 10.1016/j.displa.2022.102355.

[53] J. Chu, L. Li, and X. Xiao, Remote Sensing Image Retrieval by Multi-Scale Attention-Based CNN and Product Quantization, vol. 2021-July, pp. 8292–8297, doi: 10.23919/CCC52363.2021.9549732.

[54] Z. Wei, K. Jin, Z. Zhang, and X. Zhou, Multi-label contrastive hashing, vol. 149, no. August 2023, p. 110239, doi: 10.1016/j.patcog.2023.110239.

[55] Y. Liu, Y.-H. Zhu, X. Song, J. Song, and D.-J. Yu, Why can deep convolutional neural networks improve protein fold recognition? A visual explanation by interpretation, vol. 22, no. 5, p. bbab001, doi: 10.1093/bib/bbab001.

[56] D. Zhu, Y. Fu, X. Zhao, X. Wang, and H. Yi, Facial Emotion Recognition Using a Novel Fusion of Convolutional Neural Network and Local Binary Pattern in Crime Investigation, vol. 2022, p. 2249417, doi: 10.1155/2022/2249417.

[57] H. Afzaal, A. A. Farooque, A. W. Schumann, N. Hussain, A. McKenzie-Gopsill, T. Esau, F. Abbas and B. Acharya, Detection of a Potato Disease (Early Blight) Using Artificial Intelligence, doi: 10.3390/rs13030411.