A privacy-preserving scheme for JPEG image retrieval based on deep learning

Bin Li\textsuperscript{1}, Shilei Ding\textsuperscript{*}, Xu Yang\textsuperscript{1}

\textsuperscript{1}State Grid Electric Power Research Institute, NANJING, Jiangsu, 211100, China
\textsuperscript{*} Shilei Ding: dingshilei@sgepri.sgcc.com.cn

Abstract. With the great increase of digital images, content-based image retrieval (CBIR) has been proposed to facilitate the usage of images. Local users need to outsource CBIR to cloud services, which leads to privacy leakage. Based on deep learning, this paper proposes a retrieval scheme of encrypted JPEG images. In this scheme, the image owner sends the corresponding encrypted image to the cloud server. The cloud server extracts the Variable-length Integer (VLI) code length from the encrypted image and uses Convolution Neural Network (CNN) to learn a value of fixed dimensions as the feature vector. We have shown through experiments that our scheme is safe, with good performance and retrieval accuracy.

1. Introduction

With the rapid development of imaging equipment, especially smart mobile devices, a large number of images are generated. The images contain various information and have been playing very important roles in all the fields of our social life. In order to find the wanted image from huge image databases, many image retrieval schemes are designed to search similar images to a query. However, excessive image data often results in large storage and computation overheads. Thus, it is attractive to outsource the image storage and image-relevant services to cloud servers which can provide on-demand storage and computing resources. In this way, the image owner need not store the image database locally while the image can be retrieved directly from the cloud server.

When it comes to outsourcing images to the cloud server, privacy is a top concern to the image owner. Images contain all kinds of sensitive information. And the cloud server is never fully secure and trustworthy to the image owner. First, hackers are more interested in the information on the cloud server, so the cloud server is always the key target of hacker attacks. For example, Sony's PlayStation server was hacked in 2011. The service was interrupted for more than a week, and personal information of 77 million accounts on the network was stolen. Next, the cloud server cannot be fully trusted as they may also be interested in the information of the image owner. A straightforward way to protect sensitive image is to encrypt the image by standard cryptographic methods. However, in such a way the possible uses of the stored images would be strongly limited together with the services provided to users, such as image retrieval.

Many privacy-preserving image retrieval schemes have been proposed, and they can be divided into two categories. The first category can be named as the feature-encryption based image retrieval scheme. In this kind of scheme, the image owner extracts features from the plaintext image. Then, the images and the extracted features are encrypted and uploaded to the server, respectively. The other category is the image-encryption based scheme. In the schemes of this category, the images are encrypted by the image owner with some specially-designed encryption methods. Then, the feature extraction together with the search operation will be outsourced to the cloud server so as to further...
relieve the burden of the image owner. Generally, the second category of scheme is the more desirable ones since it outsourced more tasks to the cloud server. But the existing schemes of the second category still have many limitations in terms of security, storage expansion and retrieval accuracy.

In this paper, we propose a privacy-preserving CBIR scheme, especially for JPEG images. The main contributions of our scheme can be summarized as follows:

- A format-compatible encryption method is proposed for JPEG images without any storage expansion. Firstly, all the Discrete Cosine Transformation (DCT) coefficients are protected by encrypting the binary code in VLI coding. Secondly, the quantization tables are also protected by the stream cipher. Finally, the positions of Discrete Cosine (DC) coefficients are further permuted. The proposed method can protect the image content without storage expansion.
- For the first time, we applied the deep learning algorithm to learn features from the intermediate data which parsed from the encrypted images. Then, the Euclidean distance between the features are calculated to measure the similarity of images. Due to the powerful learning capability of neural network, the retrieval accuracy of the proposed scheme is excellent.

2. Related work

CBIR aims to find some similar images to a query image in real life[1]. In the privacy-preserving CBIR scheme, both the image and its features should be properly protected. The existing privacy-preserving CBIR scheme can be divided into two categories, i.e., feature-encryption based image retrieval schemes and the image-encryption based image retrieval schemes.

In the feature-encryption based schemes, the image owner needs to extract features from the original image at first. Then, the images and the features are encrypted respectively. Weng et al. [2, 3] proposed the privacy-preserving schemes by using locally sensitive hash (LSH). During the search process, the image owner omits or encrypts a part of the hash bits, and the remaining bits are submitted to the cloud server to search the similar images. Finally, the returned similar images are refined with the complete hash bits. Huang et al. [4] proposed an interactive image retrieval scheme which keeps user’s search intention unrevealed. Firstly, the query user provides a private query to get a private feedback from the CBIR service provider. Next, a refined set of results can be got by conducting a local retrieval at the user side. The private query is generated by adjusting the percentage of variance covered in principal component analysis (PCA) process. Yan et al. [5] proposed using Intel Software Guard Extension (SGX) for secure similar image search. The search operation is conducted in the so-called enclave where all the data are kept secure. It is too expensive to put all of these kinds of computation the enclaves, which is also revealed in their experiments.

In feature-encryption based schemes, the image owner needs to outsource the task of extracting image features to the cloud server. In image-encryption based schemes, the image owner only needs to encrypt the images. The key point is how to encrypt images and make sure that useful features can be extracted from such encrypted images.

To solve the problem of encryption breaking the dependency between pixels, Cheng et al. [6] proposed a privacy-preserving CBIR especially for JPEG images. In this scheme, the DCT coefficients are permuted at the same location in different blocks at the image owner side. Then, histogram of DCT coefficients are calculated to measure the similarity between images. Liang et al. [7] proposed a JPEG encrypted image retrieval scheme based on Huffman coding. They scrambled the AC coefficients. DC coefficient, VLI code, and quantization table are encrypted by a stream cipher. The similarity between images is measured by the Huffman coding histogram [8].

In this paper, we propose a secure outsourcing CBIR scheme for JPEG images. Deep learning is adopt to learn useful features for the image retrieval. An improved retrieval accuracy is achieved. In addition, the image encryption method in our scheme makes no size expansion of the image file. Except the image encryption, all of other tasks, i.e., feature extraction and search operation, are outsourced to the cloud server.
3. The proposed CBIR scheme

3.1. Overview of the proposed scheme

The algorithm of the proposed scheme is mainly performed by two different entities, namely the image owner and the cloud server. In addition to key generation and encrypted images, the image owner is also responsible for submitting query requests and decrypting similar images finally. The cloud server performs index construction and search similar images. The main process of the proposed scheme is shown in Figure 1.

Figure 1. Main process of the proposed scheme

In the initial phase, the image owner owns the image database \( I = \{I_i\}_{i=1}^n \) and the corresponding identification set \( ID = \{ID_i\}_{i=1}^n \). It generates the key \( K \) by a key generation method. Then, the image owner encrypts the JPEG image database to generate an encrypted JPEG image database \( C = \{C_i\}_{i=1}^n \), and uploads the encrypted image database \( C \) to the cloud server. After receiving the encrypted image database \( C \), the cloud server builds the index \( Idx \).

In the query phase, the query image can be encrypted in the same way as the image database by the image owner. Then, a query request will be generated and submitted to the cloud server. The cloud server runs the search program and calculates the similarity between the query image and the image in the database \( C \). After receiving the search result from the cloud server, the image owner eventually decrypts these similar images.

3.2. Privacy-preserving CBIR scheme

Privacy-preserving CBIR scheme mainly includes three parts: image encryption, image feature extraction and image retrieval. Next, we will introduce the main steps of our scheme in detail.

3.2.1. JPEG image encryption

Our encryption method is based on the bit-stream of a JPEG image. The image owner first reads the JPEG file information, and then uses the image data to parse the DCT coefficient entropy coding bit sequence of all components in the JPEG image. Then, the image owner can parse the bit sequence of all DCT coefficients. After parsing the JPEG bit-stream, the image can be encrypted by changing the value of the bit sequence.

Through the above steps, the encryption bit-stream of a JPEG image can be completed, and then the file is written byte by byte to obtain the encrypted JPEG image. These steps can protect the image content and retain some encrypted statistical features for image retrieval.

3.2.2. Image feature extraction

For image retrieval, feature extraction must be performed before retrieval. Here, we propose a method based on deep learning to extract features of encrypted JPEG images. The theoretical basis is that similar JPEG images have similarities in DCT coefficients. The detailed process is as follows:

First, the cloud server extracts the header file information and the Huffman coding of the DCT coefficients by parsing the encrypted JPEG bit-stream. If the coefficient is a DC coefficient or zero-value AC coefficient, its corresponding matrix element will be set to zero. Otherwise, the matrix element value of the corresponding position can be replaced by the group index of the non-zero AC coefficient. Then, we extract the features from matrices, for which the deep learning introduced
Deep learning proposes a method to automatic learn features, thereby reducing the shortcomings caused by manual features.

The inspiration comes from the construction of CNN architecture, as shown in Figure 2. The detailed process of our CNN will be described later. Then, the value of output \( k \) is artificially set, which also represents the feature dimension of the image. So that, each image can be represented by a feature vector \( f = \{ f_i \}^k_{i=1} \), and a linear search can be performed. The image identification and image feature vectors can be combined into an index table.

3.2.3. Image retrieval

The image owner will encrypt the JPEG image and then upload it to the cloud server if he wants to query similar images. After receiving the encrypted query JPEG image, the cloud server uses the same CNN model and parameters to calculate the corresponding \( k \)-dimensional feature vector as the query feature vector \( f_q = \{ f_i \}^k_{i=1} \). The cloud server uses \( f_q \) and the index table at this stage. It can directly calculate the distance between \( f_q \) and the feature vector \( f \) in the index table. In this process, each image feature of all encrypted images will be traversed and the distance between the feature vectors and \( f_q \) will be calculated. The cloud server arranges these images by distance and selects the most similar \( \theta \) image to return to the image owner while stopping the retrieval service.

The returned images are all encrypted images, and the image owner needs to decrypt these images by using the key after receiving the encrypted image. The decryption process is exactly the opposite of the encryption process, which we will not describe in detail here.

4. Experimental results

This section is intended to introduce the performance of our method primarily. We use Matlab 2016a to run the program. The device is configured as Intel Core(TM) i7-7700HQ, CPU 2.80 GHz, 8 GB RAM, Win10 system. We used the Corel Image Database as an experimental database, which contains 100 images in every category, and each image with a size of 384*256 or 256*384.

In our experiments, these 1000 images are encrypted by our proposed encryption method, and these corresponding ciphertext images were used as query images. In order to evaluate the search performance of this program, we use a common search standard: precision and recall. The precision is the probability that the returned image is similar to the query image. The recall refers to the probability that an image in the image database, which is similar to the query image, returns to the image owner.

4.1. The effectiveness of image encryption

In our scenario, the JPEG image is encoded to form the JPEG bit-stream. Then, the encrypted JPEG bit-stream can be formed by stream cipher encryption and permutation encryption. In addition, the image information becomes chaotic and disordered after the encrypted JPEG image is generated. As shown in Figure 3, the image information is greatly changed and the encrypted image is very different from the original image, and the image information is well protected.
4.2. The precision of linear retrieval

In our experiment, for the image database used, first, calculate the P-R curve of the proposed method, and then calculate mAP. The retrieval accuracy of the proposed scheme mainly depends on the dimension k of the CNN output layer, that is, the dimension of the feature vector. We set different parameters to calculate the effectiveness of our solution, as shown in Table 1. The retrieval accuracy corresponding to different parameters is shown in Figure 4.

| k  | 32  | 64  | 128 | 256 |
|----|-----|-----|-----|-----|
| mAP| 0.4445 | 0.4276 | 0.4754 | 0.4528 |

The experimental results show that the change of the parameter k affects the final retrieval accuracy, and the retrieval accuracy is the highest when k=128 among several different parameters. The results show that the dimension of the feature vector and the accuracy of the retrieval are nonlinear.

4.3. The efficiency of linear retrieval

In the proposed scheme, for the JPEG bit-stream, the DCT coefficients and the quantization tables are encrypted by stream cipher encryption and the DC coefficients are scrambled and encrypted. Image encryption is fast, and the proposed image encryption method is more efficient than some other encryption methods. The cloud server trains the image database using CNN, and each image outputs a value of a fixed dimension as the feature vector. The time required is related to the specific parameters used in the network. Linear indexes are used in our scheme, so the cloud server needs to search the index table of the entire image database to find the most similar θ images. In addition, the change in the dimension k of the feature vector also affects the time consumption. The time consumption of
different operations with different parameters $k$ is shown in Table 2. Although our retrieval scheme involves image encryption and deep learning, the time consumption of our scheme is not high.

| $k$  | Feature (s) | Search (s) |
|------|-------------|------------|
| 32   | 23.395      | 0.229      |
| 64   | 24.053      | 0.252      |
| 128  | 24.793      | 0.306      |
| 256  | 26.502      | 0.328      |

5. Conclusion
This paper proposes a new privacy-preserving encrypted JPEG image retrieval scheme that outsources the main computational overhead to the cloud server, which further reduces the burden on the image owner. The scheme protects the image content by using stream cipher encryption and scrambling encryption. It extracts the VLI coding from the encrypted image to construct matrices and uses the CNN frame to generate an output value as the feature vector representing each image on the basis of deep learning. The similarity between images is measured by the Euclidean distance between the feature vectors. Experiments have shown that our solutions are proven to be safe, having good performance and retrieval accuracy. In future work, we consider building more efficient network to improve retrieval accuracy.

Acknowledgments
This work was supported by State Grid project “On-site operation control of substation”. The authors would like to thank Kun Hu for excellent technical support and Professor Chunyong Yin for critically reviewing the manuscript.

References
[1] Rui, Y., Huang, T., Chang, S.-F. (1999) Image retrieval: Current techniques, promising directions, and open issues. Journal of visual communication and image representation., 10(1): 39-62.
[2] Weng, L., Amsaleg, L., Morton, A., Marchand-Maillet, S. (2015) A privacy-preserving framework for large-scale content-based information retrieval. IEEE Transactions on Information Forensics and Security., 10(1): 152-167.
[3] Weng, L., Amsaleg, L., Furon, T. (2016) Privacy-preserving outsourced media search. IEEE Transactions on Knowledge and Data Engineering., 28(10): 2738-2751.
[4] Huang, Y., Zhang, J., Pan, L., Xiang, Y. (2020) Privacy protection in interactive content based image retrieval. IEEE Transactions on Dependable and Secure Computing., 17(3): 595-607.
[5] Yan, H., Chen, Z., Jia, C. (2019) SSIR: Secure similarity image retrieval in IoT. Information Sciences., 479: 153-163.
[6] Zhang, X., Cheng, H. (2014) Histogram-based retrieval for encrypted JPEG images. In: 2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP), Xi’an. pp. 446-449.
[7] Liang, H., Zhang, X., Cheng, H. (2019) Huffman-code based retrieval for encrypted JPEG images. Journal of Visual Communication and Image Representation., 61:149-156.
[8] Lin, K., Yang, H. F., Hsiao, J. H., Chen, C. S. (2015) Deep learning of binary hash codes for fast image retrieval. In: IEEE Conference on Computer Vision & Pattern Recognition Workshops., Boston, MA. pp. 27-35.