

# A NOVEL METHODOLOGY FOR SECURE DEDUPLICATION OF IMAGE DATA IN CLOUD COMPUTING USING COMPRESSIVE SENSING AND RANDOM PIXEL EXCHANGING

<sup>1</sup>PRATHAP ABBAREDDY, <sup>2</sup>SREEDHAR BHUKYA, <sup>3</sup>CHANDRAMOULI NARSINGOJU,  
<sup>4</sup>B NARSIMHULU

<sup>1</sup>Staff Software Engineer

<sup>2</sup>Professor, Department of CSE, Sreenidhi Institute of Science and Technology, Hyderabad.

<sup>3</sup>Assistant Professor, Department of CSE, Vaageswari College of engineering, Karimnagar

<sup>4</sup>Assistant Professor, Department of CSE, SR University, Warangal

E-mail: <sup>1</sup>prathap.abbareddy@gmail.com, <sup>2</sup>sreedharb@sreenidhi.edu.in, <sup>3</sup>cmnarsingoju@gmail.com  
<sup>4</sup>bnreddy25@gmail.com

## ABSTRACT

With cloud computing technology, managing multimedia content became easier for organizations in general and commercial content owners. However, duplication of data causes unnecessary burden over cloud resources. At the same time, there is security to be provided to multimedia content so as to ensure data integrity. In this paper we considered these two concerns while proposing a novel methodology for secure deduplication of image data in cloud computing. We proposed a framework known as Image Security and Deduplication Framework (ISDF) which exploits compressive sensing and random pixel exchanging for security and a deduplication mechanism for getting rid of duplicate images while storing in cloud resources. Compressive sensing is a signal processing technique used to leverage image processing and image security. We proposed an algorithm named Deduplication and Secure Image Storage and Retrieval in Cloud (DSISRC). This algorithm exploits deduplication mechanisms and security mechanisms for efficient management of image content in the cloud. Besides the deduplication process benefits Cloud Service Provider (CSP) with optimal storage and processing leading to conservation of resources. A benchmark dataset is used for our empirical study. Experimental results revealed that the proposed algorithm performs well in terms of image security and deduplication.

**Keywords** – *Secure Deduplication, Compressive Sensing, Data Anonymization, Cloud Computing, Image Security*

## 1. INTRODUCTION

Multimedia content owners are increasingly relying on cloud resources for managing their digital objects. Cloud computing bestows plenty of advantages to content owners. They include affordable storage services, scalability, availability and on-demand service provisioning. However, when cloud resources are being used by service consumers, there might be data redundancy which causes severe performance issues to cloud service providers (CSP). From a CSP point of view deduplication is indispensable to optimize resource consumption. From the consumer point of view, it is important to secure image content when it is in transit and at rest. Deduplication, as shown

in Figure 1, is the process of eliminating duplicate images stored in cloud resources. Many researchers, as explored in [6], [10] and [14]. Deduplication helps service providers to optimize resource consumption and serve consumers with increased efficiency. This process, however, does not cause any problems to consumers as it manages without losing data of consumers.

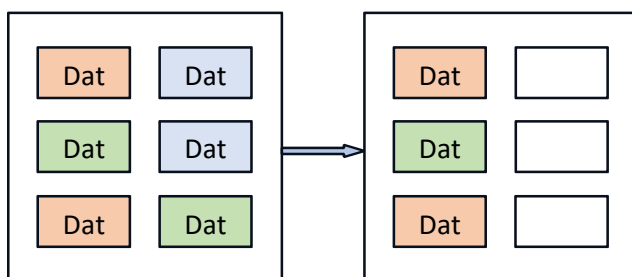


Figure 1: Illustrates how deduplication optimizes storage space in cloud

Different techniques were found in literature for deduplication and image security. Yuan et al. [6] proposed a secure deduplication scheme with efficient re-encryption using CAONT and Bloom filter-based location selection, resistant to stub-reserved attacks. Ebinazer et al. [9] enhanced Symmetric Key Encryption Algorithm, improves secure cloud storage deduplication. Athira [17] proposed secure Data Deduplication (DD) and Data Portability (DP) for distributed cloud using innovative algorithms, demonstrating effectiveness. Gao et al. [19] proposed deduplicated data integrity auditing scheme ensures data integrity, supports ownership modification, and dynamic access control in cloud storage. Some researchers also used blockchain for image security and deduplication [4], [27], [30]. Chaudhari et al. [4] addressed data handling challenges in cloud storage, emphasizing image storage. Yan et al. [27] explored a blockchain-based scheme with fair arbitration, public auditing, and secure deduplication addressing challenges in cloud storage, ensuring integrity, efficiency, and effectiveness. Pani et al. [30], studied the security aspects of networks that are used for electronic currencies. issues. From the literature review, it is understood that there is a need for novel methodology that considers both image security, lightweight cryptography and efficient deduplication. Towards this end, our contributions in this paper are as follows.

1. We proposed a framework known as Image Security and Deduplication Framework (ISDF) which exploits compressive sensing and random pixel exchanging for security and a deduplication mechanism for getting rid of duplicate images while storing in cloud resources.
2. We proposed an algorithm named Deduplication and Secure Image

Storage and Retrieval in Cloud (DSISRC).

3. We built an application to realize the proposed framework and underlying algorithm besides evaluating its performance. Our empirical results revealed that our framework is useful for secure image management in the cloud with deduplication enabled.

The remainder of the paper is structured as follows. Section 2 reviews literature on existing methods used for image security and deduplication. Section 3 presents our methodology for leveraging deduplication of images and security in the cloud. Section 4 presents our experimental results. Section 5 concludes our work besides giving scope for future work.

## 2. RELATED WORK

Wang et al. [1] proposed a secure cloud-based image protection scheme with convergent encryption and deep learning for efficient deduplication and near-duplicate detection. Nanda et al. [2] proposed a secure deduplication scheme for cloud storage, utilizing TEE for privileged user-based convergent encryption, ensuring confidentiality and efficiency. Jung et al. [3] introduced a secure protocol for single-server nearly-identical deduplication using secure LSH, proven secure against malicious adversaries, with practical efficiency. Chaudhari et al. [4] addressed data handling challenges in cloud storage, emphasizing image storage issues. Integrating data deduplication and Blockchain enhances efficiency and security. Latha et al. [5] proposed ESCDIP algorithm that improves file and content deduplication in the cloud, offering enhanced security, reducing upload/download time, and communication cost. Yuan et al. [6] proposed a secure deduplication scheme with efficient re-encryption using CAONT and Bloom filter-based location selection, resistant to stub-reserved attacks.

Raman et al. [7] introduced a secure data deduplication scheme addressing fault tolerance, efficient key management, and data confidentiality through encryption, distribution, and proof of ownership. Bhanu et al. [8] introduced SDD-RT-BF, a secure deduplication model in cloud computing, utilizing radix trie and Bloom filter for efficiency. Experimental results show its superiority over alternative

models. Ebinazer et al. [9] enhanced Symmetric Key Encryption Algorithm, improves secure cloud storage deduplication. Utilizes block-level deduplication, CE, and SMOA for efficiency and effectiveness over SKEA. Gaur et al. [10] proposed QuickDedup, a novel VM deduplication approach, which reduces time, metadata, and hash computations, outperforming traditional methods by up to 96%. Zhang et al. [11] proposed a secure deduplication scheme, utilizing proof of ownership and key-sharing, efficiently addressing convergent encryption vulnerabilities and key management issues. Yang et al. [13] The SMACD scheme in mobile cloud computing ensures media privacy through attribute-based encryption, multi-level access policies, and deduplication with low costs.

Miao et al. [14] proposed secure deduplication scheme allows efficient user revocation through a multi-user updatable encryption, minimizing communication and computation costs. Xiao et al. [15] proposed a secure medical data service with application-aware deduplication, leveraging fog-to-multi-cloud storage for efficiency, confidentiality, and fault tolerance. Singh et al. [16] EABAC-SD improves Cui et al.'s scheme for secure big data storage in the cloud, enhancing ownership management and reducing overhead. Athira [17] proposed secure Data Deduplication (DD) and Data Portability (DP) for distributed cloud using innovative algorithms, demonstrating effectiveness. Tripathy et al. [18] proposed SEDS is a server-aided data deduplication scheme for cloud storage, providing fixed-size cipher texts and efficient duplication checks across key servers. Gao et al. [19] proposed deduplicated data integrity auditing scheme ensures data integrity, supports ownership modification, and dynamic access control in cloud storage. Saini et al. [20] introduced fog-assisted cluster-based IIoT with task allocation, secure deduplication, and enhanced performance in latency, security, and energy consumption. Chitra et al. [21] proposed a secure cloud deduplication scheme using ECC-CRT for key generation, ensuring data security and efficient storage. Prakash et al. [22] focused on IoT and Cloud Computing synergy for cancer prediction, emphasizing health data encryption and efficient cloud storage. Smys et al. [23] proposed method introduces a mixed-mode analytical architecture with three-level mapping to address data deduplication

efficiently in cloud storage, optimizing storage space.

*Table 1: Summary of existing deduplication methods*

Reference	Methods	Advantages	Limitations
[3]	Secure Deduplication, Fuzzy Deduplication and Secure Locality Sensitive Hashing	In presence of adversaries, the system is able to provide security and deduplication.	It is tested under a single-server environment and needs to be evaluated in a multi-server environment.
[4]	Deduplication and Blockchain	Blockchain based deduplication provides inherent security bestowed by the technology.	Consensus algorithm needs to be improved.
[5]	Content deduplication and security	Enhanced security and storage efficiency.	It does not work for encrypted data and privacy is not yet addressed.
[8]	Radix trie based secure deduplication	Tag consistency, client side support and fault tolerance.	Performance needs to be improved with queuing methods and lightweight primitives of cryptography.
[10]	VM deduplication in cloud	Faster and efficient storage management in the cloud.	Need to be enhanced to leverage security.
[12]	Key sharing based approach in cloud	Security and efficiency	Depends on trusted entities and block level deduplication efficiency needs improvement.
[16]	Attribute based approach for deduplication	Consistency and data privacy	To be improved towards improving efficiency in deduplication
[20]	Fog assisted approach for secure deduplication	Efficient for IoT environments	To be evaluated further in healthcare

			IoT applications.
[21]	ECC based deduplication	Gets rid of malicious downloads and uploads	In the future, it needs further enhancement for efficiency.
[27]	Secure deduplication using blockchain and fair arbitration	Eliminates random masking in the process of data auditing and reduces computational complexity.	Needs improvement to work in presence of dynamically increasing users

Malleswari et al. [24] focused on research centres on live VM migration in cloud computing, using adaptive deduplication for VM disk image files, reducing storage and migration time. Conti et al. [25] observed that the cloud storage providers use deduplication to save space, but attackers can exploit it for privacy breaches. ZEUS framework introduces privacy-aware protocols, maintaining efficiency and addressing privacy concerns. Wei et al. [26] focused on efficient data deduplication in the cloud, resistant to side-channel attacks, is achieved through a fog computing-based system with improved trade-offs. Yan et al. [27] explored a blockchain-based scheme with fair arbitration, public auditing, and secure deduplication addressing challenges in cloud storage, ensuring integrity, efficiency, and effectiveness. Attigeri et al. [28] focused on a healthcare EMR processing system that uses TF-IDF, topic modelling, and KNN for classification, optimizes storage with deduplication, and secures data using DNA encryption on Hadoop. Kharade et al. [29] observed that IoT distributed computing provides on-demand processing. The proposed data deduplication optimizes caching efficiency, ensuring scalability and performance in systems. Pani et al. [30], studied the security aspects of networks that are used for electronic currencies. Table 1 shows a summary of related works. From the literature, it is observed that there is a need for a system which is capable of efficient security and also deduplication combined for protecting images in the cloud.

3. PROPOSED METHODOLOGY

Our methodology is two-fold considering image security and deduplication. Our framework, shown in Figure 1, is known as Image Security and Deduplication Framework (ISDF). It

exploits compressive sensing based encryption and also deduplication.

3.1 The Framework

An overview of the work done to improve image security when it is outsourced to the public cloud is provided by the framework built as shown in Figure 2. Multimedia content producers are progressively outsourcing their multimedia items to the cloud and sharing them with other users thanks to cloud computing technology. Since the cloud is an untrusted domain, it obviously requires a greater level of security. Within the suggested structure, the provided image is divided into four blocks. After that, it generates measurements and takes into account random pixel exchange before outsourcing the image to a public cloud.

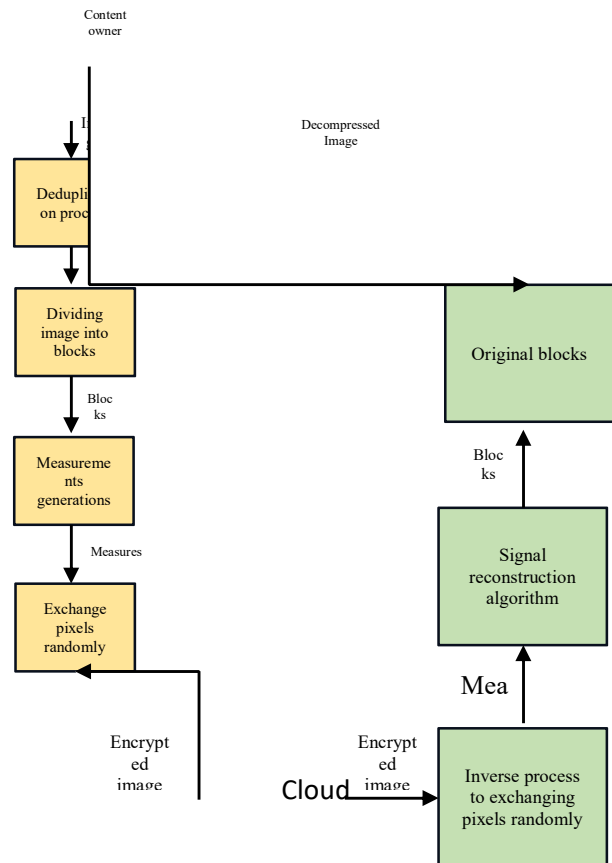


Figure 2: Proposed framework named Image Security and Deduplication Framework (ISDF) for secure and efficient image deduplication in cloud

Simultaneous encryption and compression rely heavily on a measurement matrix. For the best compression and encryption procedure to maximize image security, the matrix is managed by the logistic map. The process of obtaining the image back can be reversed once it has been outsourced to a public cloud. A random pixel exchange inverse method is used to an image that was obtained from a public cloud. It produces measures as a result, and to produce the blocks, they are then run through the signal reconstruction algorithm described in [21]. Combining the blocks yields back to the content source the original image.

### 3.2 Deduplication Method

If the image content of two tags is the same, then they are considered to be the same according to the tag correctness property [34]. An image deduplication validation phase must be carried out prior to the upload of photos in order to lower the percentage of duplicate images in the cloud. By uploading a hash value of the image content to the cloud, the user can carry out deduplication checks in the deduplication system [36]. In other words,  $H_1(I)$  can be sent to the cloud by the user to do a deduplication check. Once the user provides the hash value of the image, the cloud may verify if the identical image has been stored on the cloud server by verifying that the following equation is true:

$$\hat{e}(T_1, G) \stackrel{?}{=} \hat{e}(G, G)^{H_1(I)}. \quad (1)$$

The user does not need to submit the image of  $I$  to the cloud if the verification is successful; if not, they must encrypt the image and upload it in accordance with the instructions under "Image Uploading."

### 3.3 Detection of Near Duplicate

Image feature extraction and comparison are typically used in image processing research to accomplish image near-duplicate detection [36], [37]. This research proposes a novel near-duplicate detection method based on bilinear pairing. Additionally, under the encrypted form, near duplicate photos are discovered. To facilitate comprehension, a threshold  $t$  is established. The duplicate matrix value number of two images is represented by the value of  $t$ . Assume that  $m^*$  ( $1 \leq i \leq m^* \leq m$ ) and  $n^*$  ( $1 \leq j \leq n^* \leq n$ ) are the duplicate numbers of the row

and column values, respectively. From this, we may obtain  $t = (m^* - i + 1) \cdot (n^* - j + 1)$ . Eq. 2 represents the near-duplicate detection equation.

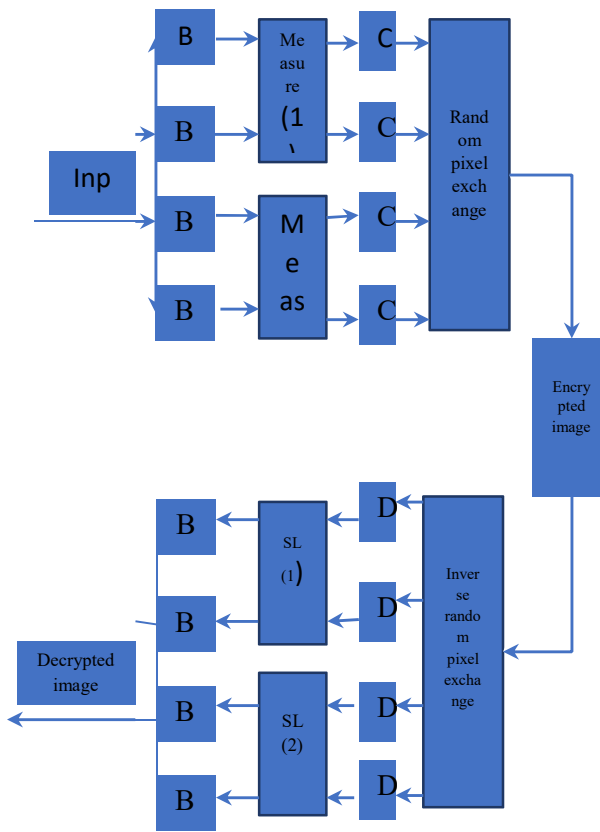
$$SP_{ij} \cdot \hat{e} \left( \prod_{i=1}^{m^*} \prod_{j=1}^{n^*} v_{ij}^{Encl_{ij}}, T_1 \right) \stackrel{?}{=} \hat{e}(T_{ij}, G), \quad (2)$$

It should be noted that the user's image parameter is calculated on the left side of Eq. 2, while the comparative image parameter is computed on the right. Assume that the two images have the same detected matrix size. The deep learning technology can be utilized by the system to rapidly and precisely select the matrix. We can see that the software may choose the image matrix based on past image matrix selection experiences by looking at the deep learning idea [38], [39]. In this technique, experiences can be generated if the system preconfigures some sample images and standards for picture matrix selection. Furthermore, the deep-learning program will continue to be executed with better image matrix selection performance.

### 3.4 Procedure for Encryption

The mechanisms depicted in Figure 3 form the foundation of the proposed method, Hybrid Image Security method (HISA). Four blocks are created from the input image. Measure 1 uses the first two blocks, whereas measure 2 takes the rest. Before being sent to the cloud, the initial measure creates two matrices that are randomly exchanged pixels. In addition, the second measure creates two matrices that, before being sent to the cloud, are randomly exchanged pixels. Following this, the process of reconstructing the original image is reversed. It first creates four matrices using the inverse random pixel exchange technique. Next, the signal reconstruction process in [21] uses the

first two matrices, followed by the second two.



```

3. IF flag=true Then
4.   Map it to existing one without saving again
5. Else
6.   Divide D into four blocks B
7. End If
8. For each two blocks in B
9.   (m1, m2)@ComputeMeasurementMatrices()
10.  Generate two matrices
11. End For
12. D@RandomPixelExchange(D, matrices)
13. D'@GenerateEncryptedImage(D)
14. Save D' to cloud
Image Retrieval
15. D'@RandomPixelExchnageReverseProcess(D')
16. (m1, m2)@DeriveMeasurementMatrices()
17. blocks@SignalReconstruction(m1, m2)
18. D@DecompressImage(blocks)
19. End
    
```

Algorithm 1 shows how the deduplication is done prior to secure storage of image in cloud. It also has the reverse process to retrieve image from cloud. The deduplication procedure is discussed earlier in Section 3 which is followed by the algorithm. If there is a duplicate image already existing, the algorithm enables the could to keep its reference instead of saving a new image again. The given image is divided into 4 blocks. Compressive sensing is employed to generate measurement matrices used to compress blocks prior to encryption. The blocks are further subjected to random pixel exchange for improving security. Finally, encrypted image is saved to cloud. In the retrieval of the image from cloud, the reverse process takes place.

Figure 3: Describe the suggested algorithm in brief

To accomplish this process, an algorithm is suggested. The definition of the proposed algorithm named Deduplication and Secure Image Storage and Retrieval in Cloud (DSISRC) is derived from the figure presented in Figure 2. In addition to using chaotic maps and randomly swapping pixels in encrypted blocks for an extra layer of security, it accepts an image as input and outputs an image after compression and encryption. Mechanisms like compressive sensing are used for concurrent compression and encryption.

**Algorithm 1:** Deduplication and Secure Image Storage and Retrieval in Cloud (DSISRC)

<p><b>Algorithm:</b> Deduplication and Secure Image Storage and Retrieval in Cloud (DSISRC)</p> <p><b>Input:</b> Image data D</p> <p><b>Output:</b> Encrypted Data D' saved to cloud</p> <ol style="list-style-type: none"> <li>Start</li> <li><b>Deduplication and Image Storage</b></li> <li>flag@PerformDeduplication(D, X) //X is existing images in cloud</li> </ol>
---

3.3 Evaluation Methodology

The effectiveness of the suggested framework is determined using a variety of assessment indicators. One common and significant statistic is correlation. An encrypted image's correlation should be close to 0, while in a high-quality image, it should be close to 1. Correlation coefficient, thus, indicates how well the suggested approach performs. Eq. 3 is used to compute it.

$$C = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum_{i=1}^N (x_i - \bar{x})^2) (\sum_{i=1}^N (y_i - \bar{y})^2)}} \tag{3}$$

Computation of the correlation coefficient can be done in three different ways. They are referred to as diagonal, vertical, and horizontal. In actuality, an encryption algorithm is key-sensitive. It does imply that a minor alteration to the key has a significant effect on the

encrypted picture. As in Eq. 4, this is quantified using mean square error (MSE).

$$MSE = \frac{1}{L \times H} \sum_{xy} [I(x, y) - D(x, y)]^2, \quad (4)$$

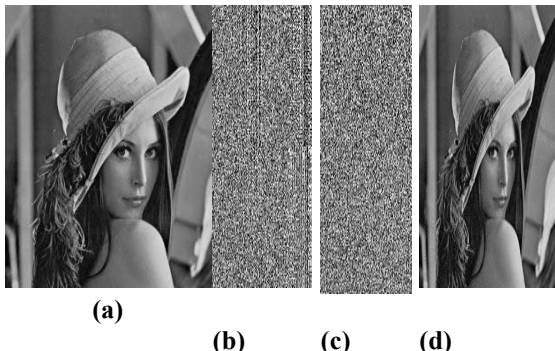
where L x H represents the total number of picture pixels. The symbols s I(x,y) and D(x,y) represent the input and output pictures, respectively. PSNR, a performance parameter that is frequently used to assess compression performance, is also used in this work. In Eq. 5, it is displayed.

$$PSNR = 10 \log \frac{255^2}{(1/N^2) \sum_i^N \sum_j^N =1 [R(i,j) - 1(i,j)]^2} \quad (5)$$

It is employed to determine the quality of decrypted pictures at different compression ratios. The quality of the suggested encryption method can be calculated by calculating PSNR.

#### 4. EXPERIMENTAL RESULTS

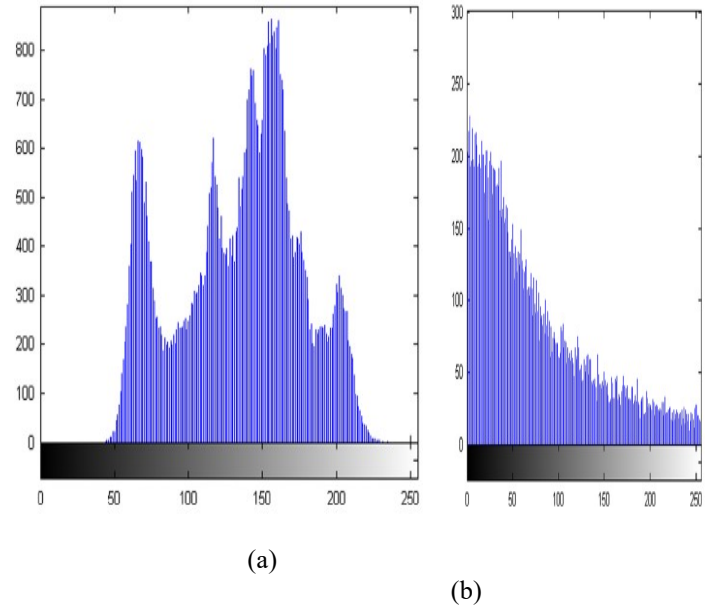
Our algorithm named Deduplication and Secure Image Storage and Retrieval in Cloud (DSISRC) is evaluated with a prototype application developing using Java language. The environment used is a PC with Windows 11 OS, Intel i3 processor 9<sup>th</sup> generation, 240 GB SSD and 6GB RAM. Experiments are made with number of images. However, some representative results are provided in this section.



**Figure 4:** Experimental results using Lena image

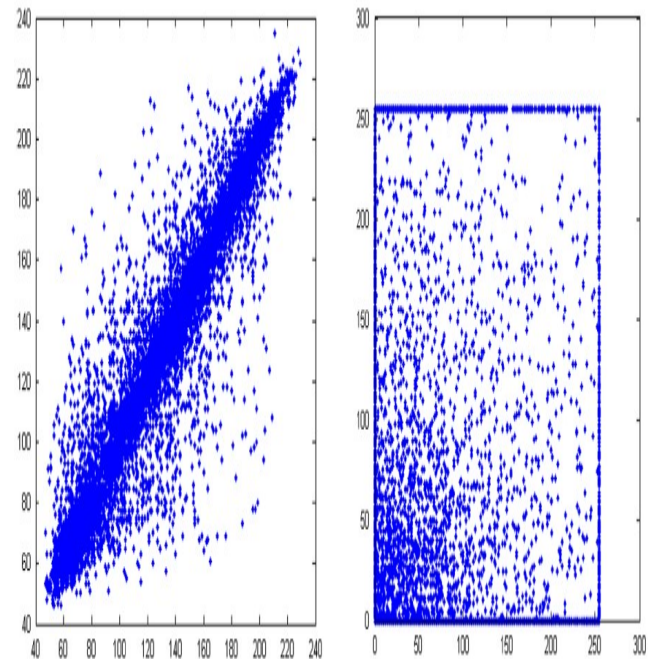
As presented in figure 4, the input image is known as Lena Figure 3(a) and it is compressed and its result is in Figure 3(b). The encrypted image as per the proposed methodology is

shown in Figure 3(c) while the decrypted image is shown in Figure 3(d). The size of the input image is 128x128. Ratio of compression used is 4/3 while key length is 2.



**Figure 5:** Lena image and its encrypted counterpart in the form of histogram

Figure 5 shows the given Lena image's histogram and its encrypted image's histograms are provided. The histograms show visible differences between the original and encrypted images.



(a)

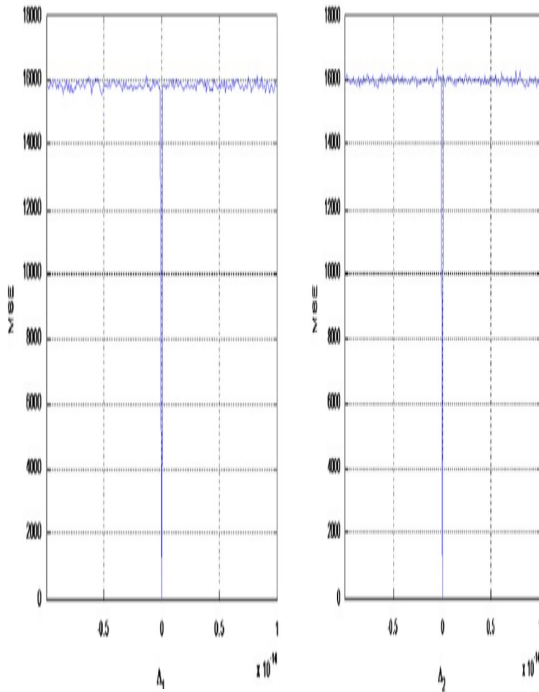
Figure 6: Lena image and its encrypted counterpart in the form of correlation distribution

Figure 6 presents pixels' correlation distribution linked to Lena image and its encrypted image. There is correlation between adjacent pixels in the original image and there is least correlation among adjacent pixels in encrypted image.

Table 2: Horizontal, vertical and diagonal correlation analysis

Input Image	Correlation Considered Horizontally	Correlation Considered Vertically	Correlation Considered Diagonally
Lena	96%	93%	91%
Lena in encrypted form	8.5%	6%	9.4%

As presented in Table 2, the horizontal, vertical and diagonal correlation of Lena image is highest while the same in encrypted form of Lena is lowest. This kind of encryption is the result of the proposed methodology. This kind of correlation does not allow attackers to gain any kind of information useful. Thus it could improve security of images significantly.



(a)

(b)

(b) Figure 7: Analysis of two decrypted images in terms of MSE curve

Figure 7 shows the visualization of analysis made on two decrypted images. It shows the key sensitivity to distortion linked to encrypted images. As the encryption methods are sensitive to keys it is important to perform MSC curve based analysis. There is higher distortion in the decrypted images reflecting the efficiency of the proposed methodology.








Input Image	Compression Ratio	Result of compression and encryption	Decrypted Image	PSNR (dB)
	4:3			35.20
	2:1			30.29
	4:1			26.39

Figure 8: Image quality and compression ratio analysis

Figure 8 shows the impact of compression ratio on the quality of the decrypted image. PSNR metric is used for performance estimation. Higher in PSNR indicates better quality. The proposed system is evaluated against attacks that are based on adding Gaussian noise. With noise addition also the proposed system is working with highly acceptable quality. With key space reduction in the proposed framework the performance is greatly enhanced. The usage of random pixel exchange method, security of the images is improved further.



## 5. DISCUSSION

The proposed framework is designed to support image security and deduplication. The security mechanism is made up of compressive sensing and pixel exchange while the deduplication is based on our methodology which identifies both exact matches and close matches. The encryption proposed in this paper is lightweight and highly secure while the deduplication is found to be robust for image data. However, it has significant limitations as described here. First, the experiments made with a limited number of samples may not be able to generalize findings. Second, it is important to consider different benchmark datasets and do experiments with diversified ratios in compression. Third, as far as security is concerned, the proposed system can be enhanced by using hybrid approaches to be comparable with Post Quantum Cryptography (PQC) standards. Fourth, the proposed system concentrated on only image data which may not be more useful unless it is extended to all kinds of data known as multimedia.

## 6. CONCLUSION AND FUTURE WORK

In this paper, we proposed a framework known as Image Security and Deduplication Framework (ISDF) which exploits compressive sensing and random pixel exchanging for security and a deduplication mechanism for getting rid of duplicate images while storing in cloud resources. Our methodology is two-fold considering image security and deduplication. Compressive sensing is a signal processing technique used to leverage image processing and image security. We proposed an algorithm named Deduplication and Secure Image Storage and Retrieval in Cloud (DSISRC). This algorithm exploits deduplication mechanisms and security mechanisms for efficient management of image content in the cloud. Besides the deduplication process benefits Cloud Service Provider (CSP) with optimal storage and processing leading to conservation of resources. A benchmark dataset is used for our empirical study. Experimental results revealed that the proposed algorithm performs well in terms of image security and deduplication. In future, we intend to improve our methodology to extend support for other multimedia objects like audio clips and videos.

## REFERENCES

- [1] Liu, Dengzhi; Shen, Jian; Wang, Anxi and Wang, Chen (2019). Secure real-time image protection scheme with near-duplicate detection in cloud computing. *Journal of Real-Time Image Processing*. <http://doi:10.1007/s11554-019-00887-6>
- [2] Fan, Yongkai; Lin, Xiaodong; Liang, Wei; Tan, Gang and Nanda, Priyadarsi (2019). A secure privacy preserving deduplication scheme for cloud computing. *Future Generation Computer Systems*, 101, 127–135. <http://doi:10.1016/j.future.2019.04.046>
- [3] Takeshita, Jonathan; Karl, Ryan and Jung, Taeho (2020). 29th International Conference on Computer Communications and Networks (ICCCN) - Secure Single-Server Nearly-Identical Image Deduplication. 1–6. <http://doi:10.1109/icccn49398.2020.9209728>
- [4] Aparna, R.; Kulkarni, Roopa G. and Chaudhari, Shilpa (2020). IEEE International Conference on Electronics, Computing and Communication Technologies (CONNECT) - Secure Deduplication for Images using Blockchain. 1–6. <http://doi:10.1109/CONECCT50063.2020.9198448>
- [5] Periasamy, J. K. and Latha, B. (2019). An enhanced secure content deduplication identification and prevention (ESCDIP) algorithm in cloud environments. *Neural Computing and Applications*. <http://doi:10.1007/s00521-019-04060-9>
- [6] Yuan, Haoran; Chen, Xiaofeng; Li, Jin; Jiang, Tao; Wang, Jianfeng and Deng, Robert (2019). Secure Cloud Data Deduplication with Efficient Re-encryption. *IEEE Transactions on Services Computing*, 1–1. <http://doi:10.1109/TSC.2019.2948007>
- [7] Singh, Priyanka; Agarwal, Nishant and Raman, Balasubramanian (2018). Secure data deduplication using secret sharing schemes over cloud. *Future Generation Computer Systems*, 88, 156–167. <http://doi:10.1016/j.future.2018.04.097>
- [8] Ebinazer, Silambarasan Elkana; Savarimuthu, Nickolas and S, Mary Saira Bhanu (2020). An efficient secure data deduplication method using radix trie with

- bloom filter (SDD-RT-BF) in cloud environment. Peer-to-Peer Networking and Applications. <http://doi:10.1007/s12083-020-00989-0>
- [9] Elkana Ebinazer, Silambarasan; Savarimuthu, Nickolas and Mary Saira Bhanu, S. (2020). ESKEA: Enhanced Symmetric Key Encryption Algorithm Based Secure Data Storage in Cloud Networks with Data Deduplication. Wireless Personal Communications. <http://doi:10.1007/s11277-020-07989-6>
- [10] Saharan, Shweta; Somani, Gaurav; Gupta, Gaurav; Verma, Robin; Gaur, Manoj Singh and Buyya, Rajkumar (2020). QuickDedup: Efficient VM deduplication in cloud computing environments. Journal of Parallel and Distributed Computing, 139, 18–31. <http://doi:10.1016/j.jpdc.2020.01.002>
- [11] Wang, Liang; Wang, Baocang; Song, Wei and Zhang, Zhili (2019). A key-sharing based secure deduplication scheme in cloud storage. Information Sciences, 504, 48–60. <http://doi:10.1016/j.ins.2019.07.058>
- [12] Wang, Liang; Wang, Baocang; Song, Wei and Zhang, Zhili (2019). A key-sharing based secure deduplication scheme in cloud storage. Information Sciences, 504, 48–60. <http://doi:10.1016/j.ins.2019.07.058>
- [13] Huang, Q., Zhang, Z., & Yang, Y. (2020). Privacy-Preserving Media Sharing with Scalable Access Control and Secure Deduplication in Mobile Cloud Computing. IEEE Transactions on Mobile Computing, 1–11. <http://doi:10.1109/tmc.2020.2970705>
- [14] Yunling Wang; Meixia Miao; Jianfeng Wang and Xuefeng Zhang; (2021). Secure deduplication with efficient user revocation in cloud storage . Computer Standards & Interfaces. <http://doi:10.1016/j.csi.2021.103523>
- [15] Yinjin Fu; Nong Xiao; Tao Chen and Jian Wang; (2021). Fog-to-MultiCloud Cooperative eHealth Data Management with Application-Aware Secure Deduplication . IEEE Transactions on Dependable and Secure Computing. <http://doi:10.1109/tdsc.2021.3086089>
- [16] Premkamal, Praveen Kumar; Pasupuleti, Syam Kumar and Singh, Abhishek Kumar; Alphonse, P. J. A. (2020). Enhanced attribute based access control with secure deduplication for big data storage in the cloud. Peer-to-Peer Networking and Applications. <http://doi:10.1007/s12083-020-00940-3>
- [17] A R Athira. (2022). Secure Data Deduplication and Data Portability in Distributed Cloud Server Using Hash Chaining and LF-WDO. Springer. 2(1), pp.1-7. <https://doi.org/10.46632/daai/2/1/2>
- [18] Nayak, Sanjeet Kumar and Tripathy, Somanath (2019). SEDS: secure and efficient server-aided data deduplication scheme for cloud storage. International Journal of Information Security. <http://doi:10.1007/s10207-019-00455-w>
- [19] Bai, Jianli; Yu, Jia and Gao, Xiang (2020). Secure auditing and deduplication for encrypted cloud data supporting ownership modification. Soft Computing. <http://doi:10.1007/s00500-019-04661-5>
- [20] Sharma, Shivi and Saini, Hemraj (2020). and MoWo in cluster-based industrial IoT (IIoT). Computer Communications, 152, 187–199. <http://doi:10.1016/j.comcom.2020.01.042>
- [21] Rasina Begum, B. and Chitra, P. (2020). ECC-CRT: An Elliptical Curve Cryptographic Encryption and Chinese Remainder Theorem based Deduplication in Cloud. Wireless Personal Communications. <http://doi:10.1007/s11277-020-07756-7>
- [22] Anuradha, M.; Jayasankar, T.; Prakash, N.B.; Sikkandar, Mohamed Yacin; Hemalakshmi, G.R.; Bharatiraja, C. and Britto, A. Sagai Francis (2020). IoT enabled Cancer Prediction System to Enhance the Authentication and Security using Cloud Computing. Microprocessors and Microsystems, 103301–. <http://doi:10.1016/j.micpro.2020.103301>
- [23] Joe, C. Vijesh; Raj, Jennifer S. and Smys, S. (2020). Mixed Mode Analytics Architecture for Data Deduplication in Wireless Personal Cloud Computing. Wireless Personal Communications. <http://doi:10.1007/s11277-020-07943-6>
- [24] TYJ, Naga Malleswari and G, Vadivu (2019). Adaptive deduplication of virtual machine images using AKKA stream to accelerate live migration process in cloud environment. Journal of Cloud

- Computing, 8(1), 3–.  
<http://doi:10.1186/s13677-019-0125-z>
- [25] Yu, Chia-Mu; Gochhayat, Sarada Prasad; Conti, Mauro; Lu, Chun-Shien (2018). Privacy Aware Data Deduplication for Side Channel in Cloud Storage. *IEEE Transactions on Cloud Computing*, 1–1. <http://doi:10.1109/TCC.2018.2794542>
- [26] Youshui Lu; Yong Qi; Saiyu Qi; Fuyou Zhang; Wei Wei; Xu Yang; Jingning Zhang and Xinpei Dong; (2021). Secure Deduplication-based Storage Systems with Resistance to Side-Channel Attacks via Fog Computing . *IEEE Sensors Journal*. <http://doi:10.1109/jсен.2021.3052782>
- [27] Yuan, Haoran; Chen, Xiaofeng; Wang, Jianfeng; Yuan, Jiaming; Yan, Hongyang and Susilo, Willy (2020). Blockchain-based public auditing and secure deduplication with fair arbitration. *Information Sciences*, 541, 409–425. <http://doi:10.1016/j.ins.2020.07.005>
- [28] A. V. USHARANI AND GIRIJA ATTIGERI. (2022). Secure EMR Classification and Deduplication Using MapReduce. *IEEE*. 10, pp.34404 - 34414. <http://DOI:10.1109/ACCESS.2022.3161439>
- [29] Ch. Prathima, Naresh Babu Muppalaneni and K. G. Kharade. (2022). Deduplication of IoT Data in Cloud Storage. *Springer*, p.147–157. [https://doi.org/10.1007/978-981-16-5090-1\\_13](https://doi.org/10.1007/978-981-16-5090-1_13)
- [30] Dileep Kumar Murala, Sandeep Kumar Panda and Santosh Kumar Sahoo. (2023). Securing Electronic Health Record System in Cloud Environment Using Blockchain Technology. *Springer*, p.89–116. [https://doi.org/10.1007/978-3-031-22835-3\\_4](https://doi.org/10.1007/978-3-031-22835-3_4)
- [31] Manikyam, N.R.H., Devi, M.S. (2021). A framework for leveraging image security in cloud with simultaneous compression and encryption using compressive sensing. *Revue d'Intelligence Artificielle*, Vol. 35, No. 1, pp. 85-91. <https://doi.org/10.18280/ria.350110>.
- [32] Mohimani H, Babaie-Zadeh M, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed l0 norm. *IEEE Trans Signal Process* 2009;57:289–301.
- [33] Mohimani H, Babaie-Zadeh M, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed l0 norm. *IEEE Trans Signal Process* 2009;57:289–301.
- [34] Bellare, M., Keelveedhi, S., Ristenpart, T.: Messagelocked encryption and secure deduplication. In: *International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 296–312 (2013)
- [35] Douceur, J.R., Adya, A., Bolosky, W.J., Simon, P., Theimer, M.: Reclaiming space from duplicate fles in a serverless distributed fle system. In: *International Conference on Distributed Computing Systems*, pp. 617–624 (2002)
- [36] Li, J., Li, J., Xie, D., Zhang, C.: Secure auditing and deduplicating data in cloud. *IEEE Trans. Comput.* 65(8), 2386–2396 (2016)
- [37] Hsu, C.Y., Lu, C.S., Pei, S.C.: Image feature extraction in encrypted domain with privacy-preserving sift. *IEEE Trans. Image Process.* 21(11), 4593–4607 (2012)
- [38] Nian, F., Li, T., Wu, X., Gao, Q., Li, F.: Efcient near duplicate image detection with a local-based binary representation. *Multimed. Tools Appl.* 75(5), 2435–2452 (2016)
- [39] Deng, L., Yu, D.: Deep learning: methods and applications. *Found. Trends Signal Process.* 7(3), 197–387 (2014)48. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 27–30 June 2016, pp. 770–778 (2016)