VISUALLY IMAGE ENCRYPTION BASED ON EFFICIENT DEEP LEARNING AUTOENCODER

Mohamed Abdelmalek[1], Anis Harhoura[1], Issam Elaloui[1],
Mahdi Madani[2] and El-Bay Bourennane[2]

[1]Faculty of Science and Technology, University of Burgundy, Dijon, France
[2]ImViA Laboratory (EA 7535), University of Burgundy, 21000 Dijon, France

ABSTRACT

This paper proposes an Artificial Intelligence (AI) model based on Convolutional Neural Network (CNN) for visual image protection during encryption and decryption processes. We used the CIFAR-10 dataset containing 60.000 color images of size 32×32 across ten classes to train and test the proposed network. Our focus lies in designing a convolutional autoencoder for image compression and reconstruction, utilizing an encoder-decoder architecture. During training, the autoencoder learns to encode essential image features into a reduced-dimensional latent space and reconstructs the image from this space. The implementation of the proposed encryption model demonstrates efficacy in preserving data integrity while reducing dimensionality. Experimental results show that the used CNN exhibits a proficient encryption process and acceptable decryption process.

KEYWORDS

Visually image protection, Deep Learning, Encryption, Encoder, Decoder, Data security, Image Compression.

1. INTRODUCTION

Today, we live in a world dominated by digital communication, wireless networks, cloud servers, connected objects, etc. Therefore, guaranteeing the security and confidentiality of user information has become a more than obligatory task. To meet these data protection requirements, several techniques have been explored for years such as encryption methods used to protect sensitive data against unauthorized access and interception, hash functions used to ensure the integrity of sensitive messages and detect any change caused by transmission errors (channel, source, etc.) or by attack. In the realm of image encryption, the quest for robust ciphering techniques has led to the exploration of traditional cryptographic methods, like chaotic systems used as pseudo number generators [1], Advanced Encryption Standard (AES) algorithm adopted in various domains [2, 3, 4], including secure communication protocols, password encryption in Wi-Fi networks, and data compression software. In the last decade, CNNs have emerged as powerful tools in the domain of image analysis and processing [5], leveraging hierarchical feature extraction through convolutional and pooling layers, CNNs excel in discerning intricate patterns and representations within image data, in the context of image encryption, CNNs offer a unique approach by learning to encode and decode images directly from their pixel values, thereby circumventing the need for explicit algorithmic rules [6].

In this work, we delve into the implementation and evaluation of a CNN-based image encryption model using a comprehensive dataset of images, particularly the CIFAR-10 dataset. The objective is the design of a new DL-based system for data privacy protection. The proposed

David C. Wyld et al. (Eds): EDUIT, BMLI, NSEC – 2024
pp. 57-64, 2024. CS & IT - CSCP 2024 DOI: 10.5121/csit.2024.141206

model utilizes a convolutional autoencoder for image compression and reconstruction, demonstrating efficacy in preserving data integrity while reducing dimensionality. Therefore, a unique model can serve as a model for two different problems of image processing applications. The architecture of a practical example of use is shown in Figure 1.

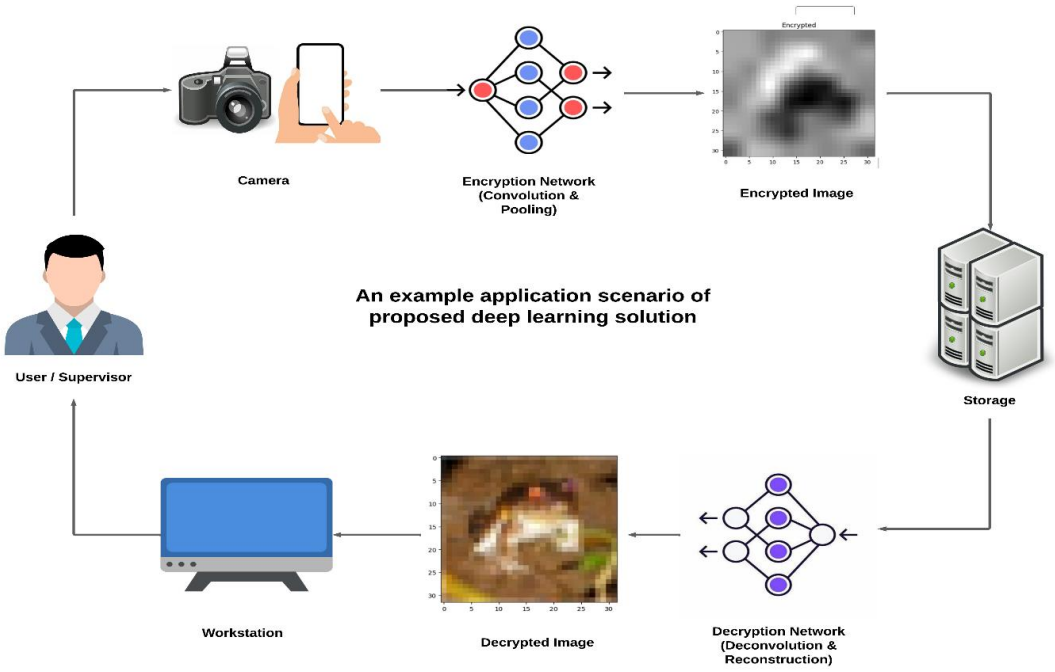

Figure 1. Deep Learning image protection model

By analyzing the performance, security level, internal architecture, and computational complexity of classic encryption techniques, we aim to provide valuable insights into the efficacy and trade-offs inherent in modern image encryption techniques [7].

The remainder of this paper is organized as follows. Section 2 discusses the whole related work, and Section 3 presents the internal architecture of the AI network and the implementation of the processing steps of the proposed DL autoencoder in both the encryption and decryption phases. Section 4 presents and discusses the experimental results. Finally, Section 5 summarizes the whole article and gives directions for our future work.

2. RELATED WORK

In the last decade, artificial intelligence has emerged in many domains of our daily life. In this section, we expose some works that used AI techniques in the area of information security principally for data privacy protection and digital data encryption.

- A Deep Learning-based Stream Cipher Generator for Medical Image Encryption and Decryption: DeepKeyGen by Yi Ding et al. [8], is a novel deep learning-based key generation network for encrypting and decrypting medical images. By employing a generative adversarial network (GAN), they aim to generate a private key, with a transformation domain guiding the learning process. DeepKeyGen seeks to learn the mapping relationship between initial images and private keys. Their evaluation on three datasets, including the Montgomery County chest X-ray dataset, the Ultrasonic Brachial

Plexus dataset, and the BraTS18 dataset, demonstrates the network's capability to achieve high-level security in key generation.

- Image to Perturbation: An Image Transformation Network for Generating Visually Protected Images for Privacy-Preserving Deep Neural Networks: Made by Hiroki Ito et al. [9], it introduces an image transformation network aimed at generating visually protected images for privacy-preserving deep neural networks (DNNs). Unlike conventional perceptual encryption methods, this network maintains image classification accuracy while exhibiting strong robustness against various attacks, including DNN-based ones. The absence of security keys further simplifies the process. Experimental validation showcases the network's capability to protect visual information while preserving high classification accuracy using ResNet and VGG classification networks. Additionally, the visually protected images demonstrate resilience against diverse attacks, affirming the efficacy of the proposed transformation network in ensuring privacy in DNN applications.
- Learnable image encryption: In recent years, many researchers have explored the existing learnable image encryption schemes and proposed new ones. Among others, we can cite Tanaka [10] who presented the state-of-the-art of privacy-preserving deep neural networks and proposed a shame based on encrypting images to will be not recognized by the human eye but still learnable by analysis with a machine, Sirichotedumrong et al. [11, 12] who presented another scheme (named as SKK scheme) using independent encryption keys unlike the basic Tanaka scheme using only one key, recently Sirichotedumrong et al. [13] proposed an image transformation scheme based on GANs, proving that the need to manage encryption keys no longer existed, and Huang [14] et al. proposed a learnable image encryption scheme that is an enhanced version of previous methods and can be used to train a great DNN model and simultaneously keep the privacy of training images.

Note that the discussed works in this section are based only on models designed to ensure image protection and data privacy. However, our proposed model utilizes a convolutional autoencoder for image compression and reconstruction, demonstrating efficacy in preserving data integrity while reducing dimensionality. The proposed model can be used on one hand as a visually encryption algorithm, and another hand as an image compression algorithm. Therefore, a unique model can serve as a model for two different problems of image processing applications.

3. PROPOSED MODEL ARCHITECTURE

In addition to the main use of CNN in computer vision (object detection, people recognition), in the last years, many works-based CNN have been published in the field of image encryption [14, 15, 16]. In this paper, we propose a visual image encryption and compression model that can be used to protect data privacy.

The designed model is based on using filters to detect patterns, edges, shapes, and colors from original images, and after several rounds of training through the backpropagation protocol, where it learns to minimize and adapt its weight and biases to approach the original and expected images. As a result, the used model allows for learning complex patterns from plaintext to generate acceptable ciphertext or to reconstruct plaintext from features of the ciphertext. Therefore, the global model known as autoencoder is formed by two main blocks, an encoder to generate ciphertext, and a decoder to reconstruct the plaintext.

The encryption process is based on implementing multiple layers, including convolutional layers followed by a RELU activation function, pooling layers, and the neural link or fully connected layers, the features of the input plain image are extracted and then used to produce the output cipher image. The details of each used layer in the architecture are given in Figure 2. This step is known as the encoder phase.

The decryption process is based on implementing the same network layers in reverse order and by replacing the convolution operations with deconvolution (transpose convolution) ones. This step is known as the decoder phase, and the details of each used layer in the architecture are given in Figure 2.

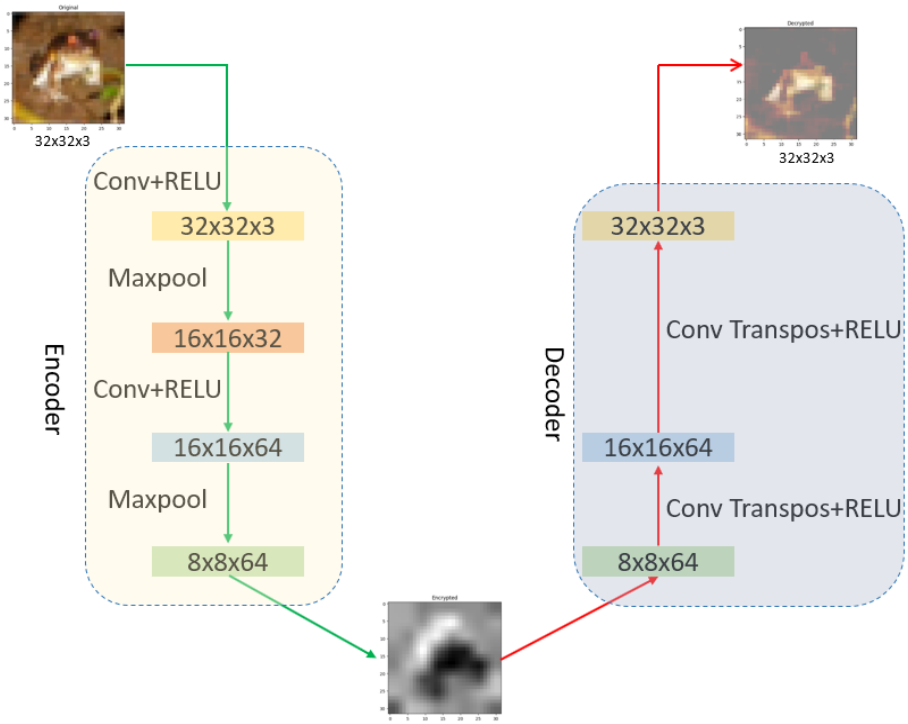

Figure 2. The autoencoder proposed architecture

After designing the neural network model firstly, it was trained and tested secondly using the CIFAR dataset. During the training step, the autoencoder learns to encrypt the important information of an image in a latent space of reduced dimension and then reconstruct the image from this latent space. To reconstruct the plain image from the visually encrypted one, we minimized the difference between both images by measuring the loss function (MSE loss) with a learning rate fixed to 0.001. As a result, the recovered image is an approximation of the original image, based on the information contained in its encoding. An example of a processed image through the autoencoder is given in Figure 2 where the visually encrypted image is generated after processing of the encoder phase, and the reconstructed image is generated after processing the decoder phase.

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we present the experimental results of the trained and we show many outputs of the proposed auto-encoder in both encoder and decoder phases.

4.1. Used CIFAR-10 Dataset

In this work, we used the CIFAR-10 database which contains 60.000 color images of size 32×32 pixels divided into 10 classes (Airplane, automobile, bird, cat, deer, dog, frog, horse, hip, truck), and every class contains 6000 images. The global space is divided into a training space which contains 50.000 images (5000 images from each class), and a test space which contains 10.000 images (1000 images from each class).

4.2. Results and Discussion

The experimental results of the trained autoencoder are shown in Figure 3.

In our analysis, we undertook several experiments to evaluate both the encryption and decryption performance of the deep learning model using 32×32 images from the known CIFAR-10 dataset. As we can see from Figure 3, it is remarkable that the encrypted images (c, f, i, l, o, and r) are visually protected and their dimension is reduced. The recovered images after the decode phase (d, g, j, m, p, and s) are noisy but still recognized compared to the original ones (b, e, h, k, n, and q). Therefore, the proposed CNN model can be useful in various image processing applications. On one hand, it can be used to visually protect images. In another hand, it can be used to image compression and reconstruction applications.

Finally, we conclude that experimental results show that the CNN model exhibits proficient encryption and acceptable decryption processes, despite some noise in the recovered images. So, it demonstrated a notable strength in image visually encryption. This discrepancy highlights the potential of deep learning models in image encryption and data privacy protection applications. Despite its shortcomings in decryption, the model's success in encryption underscores its promise in data security. Further refinement of decryption capabilities within deep learning models is necessary to fully leverage their potential in robust encryption tasks.

5. CONCLUSIONS

In this study, we conducted a comprehensive analysis of image visually protection techniques, focusing on using a Convolutional Neural Networks (CNNs) model. Utilizing the CIFAR-10 dataset comprising 60.000 color images of size 32×32 across ten classes, we implemented an autoencoder to encode (visually encrypted and dimension reduced) and decode (original image recovered) images. Experimental results show that the CNN model exhibits proficient encryption and acceptable decryption processes, despite some noise in the recovered images. However, the propped deep learning model promises since the objective was to have an output that is not understandable by humans but feels noisy in decryption capabilities.

Despite its shortcomings in decryption, the deep learning model excelled in image encryption, highlighting its potential in data security applications.

In our future works, to address the limitation of the decryption phase, one potential solution could be to implement a more advanced architecture such as a U-Net, which has shown effectiveness in various image processing tasks. Additionally, we explore implementing the proposed model on a hardware FPGA-based platform to evaluate its performance and its suitability for real-time applications.

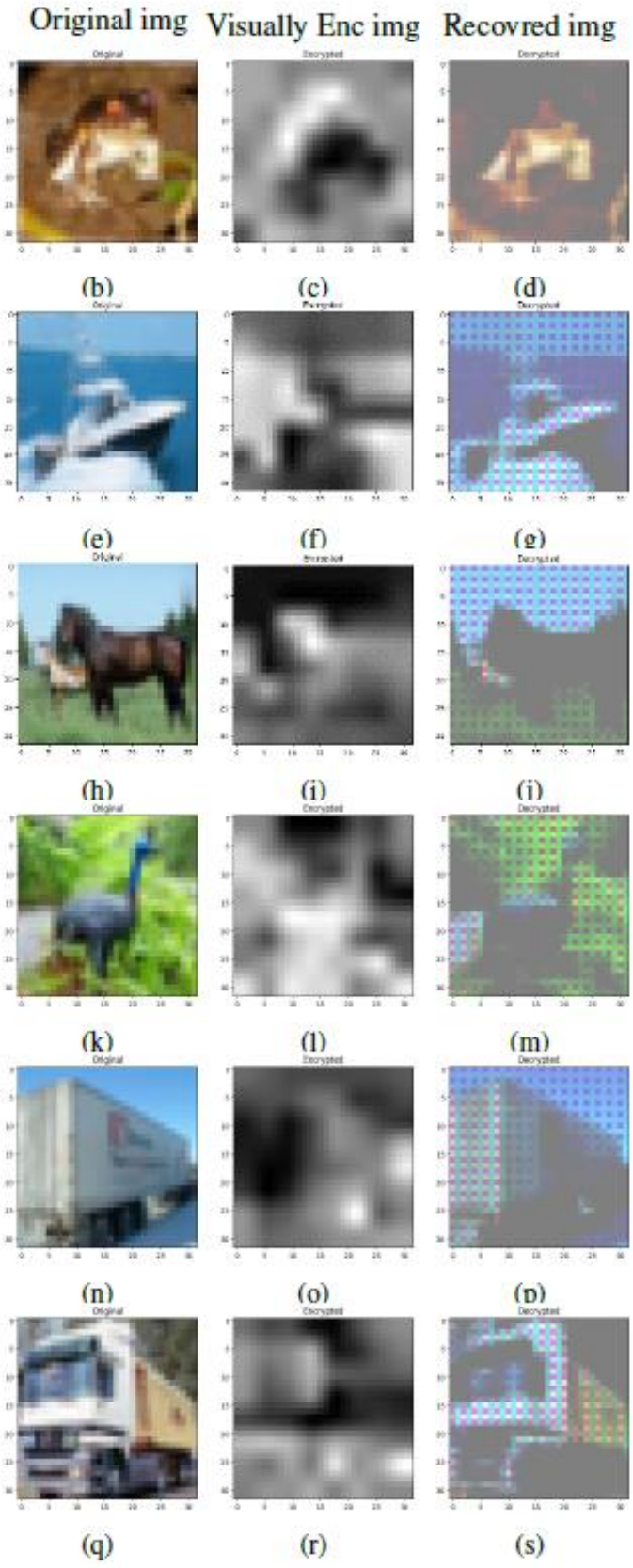

Figure 3. Experimental results for encrypted end decrypted images

REFERENCES

[1] Mahdi Madani, Safwan El Assad, F. Dridi. and Lozi, R. (2023). Enhanced design and hardware implementation of a chaos-based block cipher for image protection. *Journal of Difference Equations and Applications*, 29(9-12):1408–1428.

[2] Rijmen, V. and Daemen, J. (2001). Advanced encryption standard. In Proceedings of federal information processing standards publications, national institute of standards and technology.

[3] Assafli, H. T. and Hashim, I. A. (2020). Security enhancement of aes-cbc and its performance evaluation using the avalanche effect. In 2020 3rd *International Conference on Engineering Technology and its Applications* (IICETA), pages 7–11.

[4] Quenaya, M., Villa-Herrera, A., Chambi Ytusaca, S., Yauri Ituccayasi, J., Velazco-Paredes, Y., and Flores-Quispe, R. (2021). Image encryption using an image pattern based on advanced encryption standard. In Velasquez-Villada, C., editor, *2021 IEEE Colombian Conference on Communications and Computing, COLCOM 2021*. Institute of Electrical and Electronics Engineers Inc.

[5] D'souza, F. and Panchal, D. (2017). Advanced encryption standard (aes) security enhancement using hybrid approach. pages 647–652, 10.1109/CCAA.2017.8229881.

[6] Yi, D., Guozheng, W., Dajiang, C., Ning, Z., Linpeng, G., Mingsheng, C., and Zhiguang, Q. (2020). Deepedn: A deep learning-based image encryption and decryption network for internet of medical things. IEEE Internet of Things Journal.

[7] Khare, N., Thakur, P., Khanna, P., and Ojha, A. (2022). Analysis of Loss Functions for Image Reconstruction Using Convolutional Autoencoder, pages 338–349.

[8] Ding, Y., Tan, F., Qin, Z., Cao, M., Choo, K.-K. R., and Qin, Z. (2022). Deepkeygen: A deep learning-based stream cipher generator for medical image encryption and decryption. *IEEE Transactions on Neural Networks and Learning Systems*, 33(9):4915–4929.

[9] Ito, H., Kinoshita, Y., Aprilpyone, M., and Kiya, H. (2021). Image to perturbation: An image transformation network for generating visually protected images for privacy-preserving deep neural networks. *IEEE Access*, 9:64629–64638.

[10] Masayuki Tanaka. Learnable image encryption. *2018 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, pages 1–2, 2018.

[11] Sirichotedumrong, W., Kinoshita, Y., Kiya, H. (2019). Pixel-based image encryption without key management for privacy-preserving deep neural networks. *IEEE Access*, 7, 177844-177855.

[12] Sirichotedumrong, W., Maekawa, T., Kinoshita, Y., Kiya, H. (2019, September). Privacy-preserving deep neural networks with pixel-based image encryption considering data augmentation in the encrypted domain. *In 2019 IEEE International Conference on Image Processing (ICIP) (pp. 674-678). IEEE*.

[13] Sirichotedumrong, W., Kiya, H. (2020). A GAN-based image transformation scheme for privacy-preserving deep neural networks. *In 2020 28th European Signal Processing Conference (EUSIPCO) (pp. 745-749). IEEE*.

[14] Q. -X. Huang, W. L. Yap, M. -Y. Chiu and H. -M. Sun (2022), "Privacy-Preserving Deep Learning With Learnable Image Encryption on Medical Images," *In IEEE Access*, vol. 10, pp. 66345-66355, 2022, doi: 10.1109/ACCESS.2022.3185206.

[15] Sooksatra, K. and Rivas, P. (2020). A review of machine learning and cryptography applications. *In 2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 591–597.

[16] Li, Q., Meng, X., Yin, Y., and Wu, H. (2021). A multi-image encryption based on sinusoidal coding frequency multiplexing and deep learning. *Sensors*, 21(18).

AUTHORS

M. Abdelmalek, A. Harhoura, and I. Elaoui are Masters (M2) students in advanced electronics systems engineering at the University of Burgundy, Dijon, France, 2023-2024.

Mahdi Madani is an associate professor at the University of Burgundy, Image et Vision Artificielle (Imvia) laboratory. His main research interests are information security, design, and hardware implementation of algorithms/architectures, deep-learning for image protection applications. He received his Ph.D. degree in Electronics Systems from the University of Lorraine in July 2018. He was temporary research and teaching associate at IUT Auxerre (2 years), and IUT Nantes (2 years).

El-Bay Bourennane is currently a Professor of Electronics with the Laboratory of Image et Vision Artificielle (ImVia), University of Burgundy, Dijon, France. His research interests include dynamic reconfigurable system, image processing, embedded systems, and FPGA design and real-time implementation.