IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

SECURING VIDEO USING DEEP NEURAL NETWORK

S. Krishna Veni[1] Dr. Jai Ruby MCA, M.Phil., PhD [2]

Department of Computer Applications, Sarah Tucker College, Tirunelveli-7.

Abstract: Video in the national defense, education, monitoring, entertainment and other fields have been widely used, so data security on the internet cannot be ignored.Video encryption protects the original video information and improves the security of video information. Researchers have done a lot of research on video encryption and put forward a lot of video encryption methods. Video encryption methods are mainly divided into complete encryption and partial encryption algorithm. In order to improve the generalization performance of video encryption and reduce the amount of data in video encryption, this paper proposes a video encryption on regions of interest (ROI) method based on Faster R-CNN by combining machine learning with information security. The method trains a Faster R-CNN model using the ROI dataset firstly, and then uses the model to extract ROI in the video. Different encryption algorithms are used to encrypt ROI and non-ROI in the video respectively. To overcome the shortcomings of encryption algorithms that can only be used for a specific coded video, a special video encryption method is proposed to encrypt the video with different video coding structure and has better generalization performance. Compared with the encryption method in the video coding process, this method considers the content information of the video fully and has better performance. It can be concluded through experiments that the encryption method in this paper has the characteristics of higher security and less calculation.

I. INTRODUCTION

In the age of information, with the rapid development of computer and Internet technologies, people can post messages and obtain information on the Internet anytime, anywhere. Large amounts of data are transmitted through the network and more than 50 million servers work on the network over the world. Zettabytes of data are produced each year, which contain a large amount of video data. Unauthorized access and the openness of the network lead to more and more serious data security problems. Especially, video security issues have become more serious and aroused more attentions. Video encryption is an effective data encryption strategy to improve video security. Video complete encryption algorithm is to encrypt the whole video data with encryption algorithm in order to achieve the purpose of protecting video information. Think of video data as a series of data streams and then encrypt the video data streams with traditional encryption algorithms. However, traditional encryption algorithms such as AES and RSA can achieve good encryption effect on text data and unformatted data .In , the author proposed a video encryption algorithm based on RSA. Because of the large quantity and strong correlation of video data, the information redundancy leads to too high complexity of video encryption and too long time consuming to meet real time encryption Request. Using image encryption algorithm to encrypt the video fully, the video is divided into a series of video frames, and then use the image encryption algorithm to encrypt every frame of video. It can reduce the amount of data to be encrypted in the video, but does not take the information redundancy between video frames into account, which results in higher encryption complexity.

Video complete encryption method does not consider the video data format and ignores the correlation between video frames, resulting in higher video encryption complexity, large amount of data to be encrypted and long time consuming. So it cannot meet the demands of real time encryption. The video partial encryption algorithm encrypts the video in the coding process. Encrypting video in the process of video encoding can reduce the amount of data to be encrypted and reduce the complexity of encryption. In , the author combines the stream cipher and the video cipher to encrypt the DCT transform coefficients. This method only encrypts the DCT transform coefficients, so it is not safe enough. In , the author encrypts motion vector difference (MVD), luma residual coefficients and chroma residual coefficients in the process of HEVC encoding. It can improve encrypted security to some extent. In in order to protect the video data, the author proposed a video encryption method based on logistic chaos mapping, which encrypts the motion vector (MVD) and DCT variation coefficient the chaotic mapping in the process of HEVC encoding. In the author proposed a video encryption method based on RGB three channel MPEG encoding, which achieved wonderful encryption effect on MPEG videos, but it has a weak generalization ability. In Mamoonaetal. proposed a video encryption method based on CABAC entropy coding and it has some limitations and can only encrypt H.264 and HEVC encoded video. The existing video encryption methods combined with video coding can reduce the complexity of encryption and increase the speed of encryption. However, they also have some limitations, and can only encrypt video in a specific encoding format, and have a weak generalization ability.

In order to overcome the shortcomings of existing video encryption algorithms, this paper proposed a region-based video encryption method, which uses the Faster R-CNN to extract ROI in video frames. It can help to reduce the amount of video encrypted data. The video encryption method can encode a variety of video encoding and has a high generalization ability.

This article is organized as follows: Section 2 introduces the Faster-R-CNN network structure, training of ROI model, and extraction of ROI. Section 3 describes the encryption algorithm and the steps to encrypt the ROI. Section 4 is mainly to analyze the effect of video encryption through experiments

II. LITERATURE SUREY

In[1], This paper serves as another important security result showing that any future design of image encryption schemes based on chaotic map should be evaluated through systematic cryptanalytic approaches which include impossible differential attack. To the best of our knowledge, this is the first impossible differential attack applied on an image encryption algorithm.

In[2], The topic of semantic segmentation has witnessed considerable progress due to the powerful features learned by convolutional neural networks (CNNs). The current leading approaches for semantic segmentation exploit shape information by extracting CNN features from masked image regions. This strategy introduces artificial boundaries on the images and may impact the quality of the extracted features. Besides, the operations on the raw image domain require to compute thousands of networks on a single image, which is time-consuming. In this paper, we propose to exploit shape information via masking convolutional features. The proposal segments (e.g., super-pixels) are treated as masks on the convolutional feature maps. The CNN features of segments are directly masked out from these maps and used to train classifiers for recognition. We further propose a joint method to handle objects and "stuff" (e.g., grass, sky, water) in the same framework. State-of-the-art results are demonstrated on benchmarks of PASCAL VOC and new PASCAL-CONTEXT, with a compelling computational speed.

In[3], In this work, we propose a saliency-inspired neural network model for detection, which predicts a set of class-agnostic bounding boxes along with a single score for each box, corresponding to its likelihood of containing any object of interest. The model naturally handles a variable number of instances for each class and allows for crossclass generalization at the highest levels of the network. We are able to obtain competitive recognition performance on VOC2007 and ILSVRC2012, while using only the top few predicted locations in each image and a small number of neural network evaluations.

In[4], An alternative-transforms-based scheme has recently been proposed to achieve perceptual encryption of video signals in which multiple transforms are designed by using different rotation angles at the final stage of the discrete cosine transforms (DCTs) butterfly flow-graph structure. More recently, it is found that a set of more efficient alternative transforms can be derived by introducing sign-flips at the same stage, which is equivalent to an extra rotation angle of π. In this paper, we generalize this sign-flipping technique by randomly embedding sign-flips into all stages of the DCTs butterfly structure so that the encryption space becomes much larger to yield a higher security. We pursue this study for H.264-compatible videos, assuming that the integer DCT of size 4×4 is used. First, we follow the separable implementation of the 4×4 2-D DCT in which different sign-flipping strategies will be employed along its horizontal and vertical dimensions. Second, we convert the 4×4 2-D DCT into a 16-point 1-D butterfly structure so that more sign-flips can be embedded at its various stages. Third, we choose different schemes to pair the node-variables in the 16-point 1-D butterfly structure, thus further enlarging the encryption space. Extensive experiments are conducted to show the performance of these improved encryption schemes and some security analyzes are also presented to confirm their persistence to various attacking strategies.

In[5], Recently, creating security in multimedia systems involving video has become one of the main needs in commercial and military usages. The most important traits of video frames are simultaneousness of frames and high volume of information. Considering these features, A5/1 and W7 stream ciphers are used for video main frames in the field of selective encryption of each DCT transform coefficient. In addition, a suitable method is expressed in order to select the encryption method of DCT transform coefficients by the stream algorithms. The simulation results of MATLAB software show that this method is suitable for multimedia security in terms of both security and execution speed of the algorithm.

In[6], Now a days digital communication is a large pool of information and so its security and privacy are very sensitive and vital characteristics of the system. As far as success of any event is concerned, the keystone is effective and safe communication. We are presenting here a novel approach wherein RSA, random DNA encryption, Huffman encoding and 2D DCT steganography melds to give a system with guaranteed three levels of security. As compared to existing methods, the new approach is found to improve the quality of steganographic system and deciphering the codes will be much more cumbersome.

In[7], In this work we present a pseudo-random Bit Generator via unidimensional multi-modal discrete dynamical systems called k-modal maps. These multi-modal maps are based on the logistic map and are useful to yield pseudo-random sequences with longer period, i.e., in order to attend the problem of periodicity. In addition, the pseudo-random sequences generated via multi-modal maps are evaluated with the statistical suite of test from NIST and satisfactory results are obtained when they are used as key stream. Furthermore, we show the impact of using these sequences in a stream cipher resulting in a better encryption quality correlated with the number of modals of the chaotic map. Finally, a statistical security analysis applied to cipher images is given. The proposed algorithm to encrypt is able to resist the chosen-plaintext attack and differential attack because the same set of encryption keys generates a different cipher image every time it is used.

In[8], A digital image encryption algorithm based on dynamic deoxyribonucleic acid coding and chaotic operations using hyper digital chaos in frequency-domain is proposed and demonstrated, where both the amplitude and phase components in frequency-domain are diffused and scrambled. The proposed encryption algorithm is evaluated through various evaluations of key parameters such as histogram uniformity, entropy, and correlation. Excellent performance of the encrypted image is achieved to resist the statistical attacks, which implies that the statistical properties of the original image are completely destroyed. In the encryption procedure, each cipher pixel is affected by all of the plain-pixels as well as cipher-pixels, due to the implementation of chaotic diffusion and scrambling operations, which increases the sensitivity of the encrypted image to the plain-text, and improves the security against any differential attacks. Moreover, due to the high sensitivity introduced by the hyper digital chaos, a huge key space is provided for the encrypted image to ensure the high security level, thus the encryption algorithm has a strong secure capability against the brute-force attacks.

In[9],In order to improve the generalization performance of video encryption and reduce the amount of data in video encryption, this paper proposes a video encryption on regions of interest (ROI) method based on Faster R-CNN by combining machine learning with information security. The method trains a Faster R-CNN model using the ROI dataset firstly, and then

uses the model to extract ROI in the video. Different encryption algorithms are used to encrypt ROI and non-ROI in the video respectively. To overcome the shortcomings of encryption algorithms that can only be used for a specific coded video, a special video encryption method is proposed to encrypt the video with different video coding structure and has better generalization performance. Compared with the encryption method in the video coding process, this method considers the content information of the video fully and has better performance. It can be concluded through experiments that the encryption method in this paper has the characteristics of higher security and less calculation.

In[10], In this paper, we propose a novel video sequences compression and encryption method combining 3D compressive sensing (3D-CS) with 3D discrete fractional random transform (3D-DFrRT). In this scheme, the original video sequences were transformed with discrete wavelet and measured by three Gaussion random matrices to achieve compression and encryption simultaneously, and then the resulting 3D image was reencrypted by Arnold transform (AT) and 3D-DFrRT. Three random circular matrices used in 3D-FrRT were constructed by sine logistic modulation map. The three-dimensional smoothed `0-norm algorithm was adopted to obtain the decrypted video sequences. Simulation results verified the good compression performance, efficiency and security of the proposed method.

III. EASE OF USE

Video data often contains a lot of information, but people tend to focus only on some of the information they are inter-tested in and ignore some of the background information. In order to meet this demand, a video encryption method based on Faster RCNN and ROI is proposed in this paper. Faster R-CNN is a multi-layer convolution neural network that achieves good results in the field of object detection and recognition. In this method, the Faster R-CNN is used to extract ROI in the video. The framework of video encryption system is shown in Fig. I. It is mainly divided into three parts: the training of the ROI model, the extraction of the ROI in the video and the encryption of the ROI. The detailed ideas of this method is as follows: Firstly, a ROI model is trained with the Faster R-CNN by using the data set of the ROI, and then the ROI in the video is extracted with the trained ROI model. Finally, the ROI is encrypted.

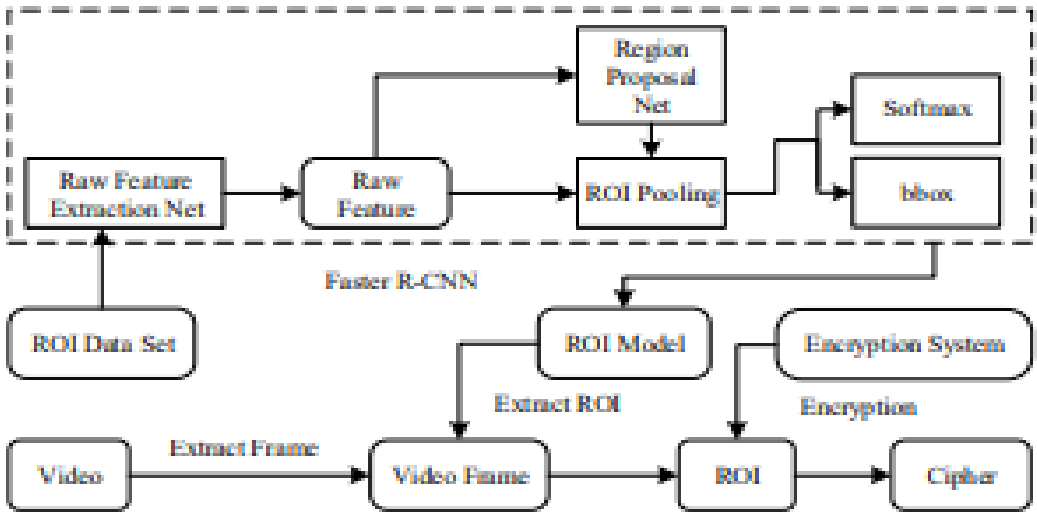

FIGURE I. THE FRAMEWORK OF VIDEO ENCRYPTION SYSTEM

A. Extraction of ROI

This section mainly introduces the extraction of the ROI in a video. In this paper, we used Faster R-CNN to extract the ROI in a video. The Faster R-CNN combines convolutional neural networks and machine learning . It uses the region proposal network instead of the selective search algorithm to generate a suggestion window, and the region proposal network and the target detection network share the convolution layer features .

The structure of Faster R-CNN is shown in Figure II. Faster R-CNN consists of the convolutional neural network, ROI Pooling, Softmax, Bounding box regression and others. Convolution neural network is mainly used for the extraction of image feature. In this paper, VGG16 is used to extract image features. Compared with other convolutionalneural networks, VGG16 has simpler structure and superior performance. The Faster RCNN with the region proposal network has better performance than the Faster R-CNN and R-CNN with the selective search algorithm in generation of the candidate box. The ROI pooling layer mainly performs the pooling operation on the candidate boxes generated by the RPN and generates a fixed-size feature map for each ROI.

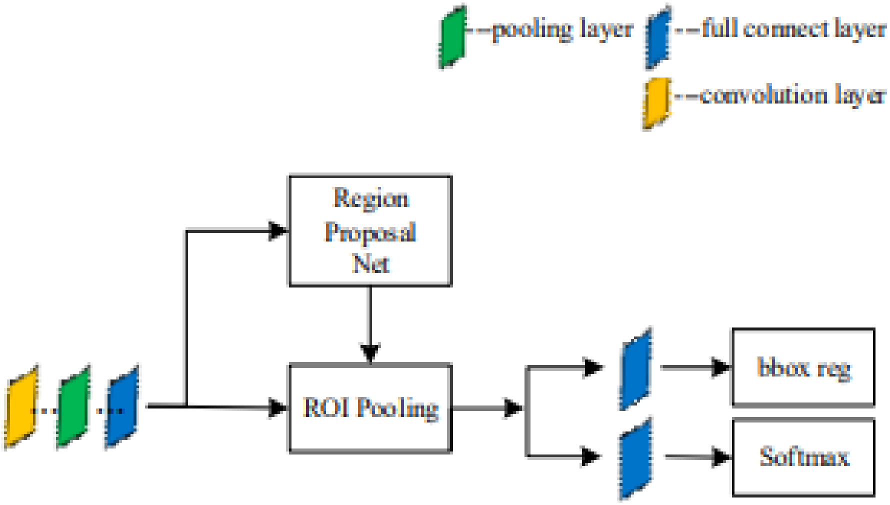

FIGURE II. THE STRUCTURE OF FASTER R-CNN

The dataset training the Faster R-CNN to generate the ROI model is from the WIDER FACE. More than 20,000 images were used to train and generate a ROI model on the GPU spending more than 10 hours. The effect of extraction of the ROI is shown in Fig. III. The model extracted the ROI in a) and extracted 5 ROI in b) accurately. It can be concluded that the ROI model in this paper is not only applicable to images containing a single ROI, but also to images containing multiple ROI.

a. Single ROI b. Multiple ROI

FIGURE III. EXTRAC ROI

B. Encryption Algorithms

After the ROI of the video is extracted using the ROI model trained, the videos are divided into the ROI and the non-ROI. In this paper, different encryption algorithms are used respectively to encrypt the non-ROI and the ROI in the video. The non-ROI of the video is encrypted by ی ఄ17ఄ domain diffusion encryption algorithm based on plaintext, while the ROI of the video is encrypted by the encryption algorithm based on hyperchos system and pixel information, which is more secure and complex. Using different encryption algorithms according to different contents of the video can improve the security of encryption and increase the difficulty of cracking and speed up encryption. The finite field, also known as the Galois field, is a field that contains only a limited number of elements. If GF(p) is a finite field, where p is a prime number, the addition is as shown in Equation 1, and the multiplication is as shown in Equation 2,where x and y are the elements in the finite field. For the multiplication in GF(P), only when pis a prime number, have the other elements in the finite field except 0 an inverse multiplication.

$$(x + y) mod\ p \quad (1)$$

$$(x \cdot y) mod\ p \quad (2)$$

TABLE I. MULTIPLICATION OPERATION OF GF(17) [8]

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	0	2	4	6	8	10	12	14	16	1	3	5	7	9	11	13	15
3	0	3	6	9	12	15	1	4	7	10	13	16	2	5	8	11	14
4	0	4	8	12	16	3	7	11	15	2	6	10	14	1	5	9	13
5	0	5	10	15	3	8	13	1	6	11	16	4	9	14	2	7	12
6	0	6	12	1	7	13	2	8	14	3	9	15	4	10	16	5	11
7	0	7	14	4	11	1	8	15	5	12	2	9	16	6	13	3	10
8	0	8	16	7	15	6	14	5	13	4	12	3	11	2	10	1	9
9	0	9	1	10	2	11	3	12	4	13	5	14	6	15	7	16	8
10	0	10	3	13	6	16	9	2	12	5	15	8	1	11	4	14	7
11	0	11	5	16	10	4	15	9	3	14	8	2	13	7	1	12	6
12	0	12	7	2	14	9	4	16	11	6	1	13	8	3	15	10	5
13	0	13	9	5	1	14	10	6	2	15	11	7	3	16	12	8	4
14	0	14	11	8	5	2	16	13	10	7	4	1	15	12	9	6	3
15	0	15	13	11	9	7	5	3	1	16	14	12	10	8	6	4	2
16	0	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1

In this paper,$GF(17)$ domain multiplication is used as operationofdiffusionencryption.Inordertoprevent0elementfrom appearing in multiplication in $GF(17)$ domain during encryption, the multiplication Operation of $GF(17)$ is shown inTable1.Forthefirsttwopixels,thedataisencryptedwithequation 3, where and the subscript idenotes the position of the dataandthesubscript$h$denotestheupperfourbitsofthedataandthe subscript $l$denotes the lower four bits of the data and the operator × is a multiplication on the $GF(17)$ field . The encryptionkey $k$is generated with the overall information of the image andthedatainformationofthepixel.Firstly,usetheupperfourbitsoftheencryptionkeytoencrypttheupperfourbitsofthedata.Secondly, usethelowerfourbitsoftheencryptionkeytoencryptthe lower four bits of the data. Finally, generate the ultimate encryptedcipher textwiththepreviousresults.

$$\begin{cases} c_{i,h} = p_{i,h} \times k_{i,h} \\ c_{i,l} = p_{i,l} \times k_{i,l} \\ c_i = c_{i,h} \cdot 16 + c_{i,l} \quad i < 3 \end{cases} \quad (3)$$

When i 3, the data is encrypted with equation 4, and the upper four bits and the lower four bits of data are encrypted respectively. Then the first two cipher texts adjacent to the data are diffused into the encrypted cipher text of the data. Finally, the final cipher text is generated with the upper four digit cipher text and the lower four digit cipher text.

$$\begin{cases} c_{i,h} = p_{i,h} \times k_{i,h} \times c_{i-1,h} \times c_{i-2,h} \\ c_{i,l} = p_{i,l} \times k_{i,l} \times c_{i-1,l} \times c_{i-2,l} \\ c_i = c_{i,h} \cdot 16 + c_{i,l} \quad i \geq 3 \end{cases} \quad (4)$$

Thedetailed procedure of decryption in this paper is as follows: Firstly, decrypt the plaintext of the first two data and then decrypt the remaining data using equation (5).

$$\begin{cases} p_{i,h} = c_{i,h} \div k_{i,h} \div p_{i-1,h} \div p_{i-2,h} \\ p_{i,l} = c_{i,l} \div k_{i,l} \div p_{i-1,l} \div p_{i-2,l} \\ p_i = p_{i,h} \cdot 16 + p_{i,l} \quad i < 3 \end{cases} \quad (5)$$

IV. EXPERIMENT AND ANALYSIS

V. CONLIUSION

This paper presented a video encryption method on ROI of a video based on Faster R-CNN. We combined machine learning with video encryption and trained an extraction of ROI model with the dataset of ROI and Faster R-CNN. The ROI in the video was extracted effectively with the trained ROI model, then the ROI and non-ROI in the video were encrypted with different encryption algorithms. The method proposed in this paper can reduce the complexity of encryption and improve the encryption speed. And it can also deal with a variety of video encoding format, so it has a high generalization performance. And the encrypted video has the advantages of low correlation, high encryption sensitivity and high security.

REFERENCES

1. Yap, Wun-She, et al. "Cryptanalysis of a New Image Alternate Encryption Algorithm Based on Chaotic Map." *Nonlinear Dynamics*, vol. 80, no. 3, May 2015, pp. 1483–91. *DOI.org (Crossref)*, https://doi.org/10.1007/s11071-015-1956-x.
2. Dai, Jifeng, et al. "Convolutional Feature Masking for Joint Object and Stuff Segmentation." *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2015, pp. 3992–4000. *DOI.org (Crossref)*, https://doi.org/10.1109/CVPR.2015.7299025.
3. Erhan, Dumitru, et al. "Scalable Object Detection Using Deep Neural Networks." *2014 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2014, pp. 2155–62. *DOI.org (Crossref)*, https://doi.org/10.1109/CVPR.2014.276.
4. Zeng, Bing, et al. "Perceptual Encryption of H.264 Videos: Embedding Sign-Flips Into the Integer-Based Transforms." *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 2, Feb. 2014, pp. 309–20. *DOI.org (Crossref)*, https://doi.org/10.1109/TIFS.2013.2293955.
5. Bahrami, Saeed, and Majid Naderi. "Encryption of Video Main Frames in the Field of DCT Transform Using A5/1 and W7 Stream Encryption Algorithms." *Arabian Journal for Science and Engineering*, vol. 39, no. 5, May 2014, pp. 4077–88. *DOI.org (Crossref)*, https://doi.org/10.1007/s13369-014-1077-8.
6. Mumthas, S., and A. Lijiya. "Transform Domain Video Steganography Using RSA, Random DNA Encryption and Huffman Encoding." *Procedia Computer Science*, vol. 115, 2017, pp. 660–66. *DOI.org (Crossref)*, https://doi.org/10.1016/j.procs.2017.09.152.
7. García-Martínez, M., and E. Campos-Cantón. "Pseudo-Random Bit Generator Based on Multi-Modal Maps." *Nonlinear Dynamics*, vol. 82, no. 4, Dec. 2015, pp. 2119–31. *DOI.org (Crossref)*, https://doi.org/10.1007/s11071-015-2303-y.
8. Guan, Mengmeng, et al. "Chaotic Image Encryption Algorithm Using Frequency-domain DNA Encoding." *IET Image Processing*, vol. 13, no. 9, July 2019, pp. 1535–39. *DOI.org (Crossref)*, https://doi.org/10.1049/iet-ipr.2019.0051.
9. Duan, Lijuan, et al. "A Novel Video Encryption Method Based on Faster R-CNN." *Proceedings of the 2018 International Conference on Computer Science, Electronics and Communication Engineering (CSECE 2018)*, Atlantis Press, 2018. *DOI.org (Crossref)*, https://doi.org/10.2991/csece-18.2018.21.
10. Alfalou, A., et al. "Simultaneous Compression and Encryption of Color Video Images." *Optics Communications*, vol. 338, Mar. 2015, pp. 371–79. *DOI.org (Crossref)*, https://doi.org/10.1016/j.optcom.2014.10.020.