Mathematics Interdisciplinary Research **x** (202x) $xx - yy$

Providing a Framework for Increasing Security in Digital Document Management Systems

Bashir Omrani Harzand, Mohammadreza Motadel ⋆

and Ali Broumandnia

Abstract

In the era of information technology and the expansion of global networks, digital document management systems play a crucial role in providing services. The use of digital instead of paper documents has become more common to reduce the administrative costs in small and large organizations. One of the most significant concerns of such systems, though, is the security of documents against alteration or distortion. There are various ways to increase the security of digital documents and this paper is aimed at providing a new framework of the combination of steganography and cryptography to increase the security of digital documents.

The proposed method first has been tested and reviewed with MATLAB 2018 software on standard images such as Lena, Baboon and Pepper. Then, it has been tested on digital documents of the Social Security Organization through the proposed conceptual model which utilizes a combination of unique image features to generate messages. The evaluation of results indicates the effectiveness of the proposed method and its applicability in digital document management systems.

Keywords: document management, image steganography, cryptography, security.

2020 Mathematics Subject Classification: 94A62, 94A60, 94A08.

How to cite this article
B. Omrani Harzand, M. Motadel and A. Broumandnia, Providing a framework for increasing security in digital document management systems, *Math. Interdisc. Res.* **x** (202x) xx-yy.

⋆Corresponding author (E-mail: moh.motadel@iauctb.ac.ir)
Academic Editor: Seyed Morteza Babamir
Received 14 March 2021, Accepted 28 June 2022
DOI: 10.22052/mir.2022.242134.1280

1. Introduction

Thanks to the development of communication platforms, information transfer in the form of digital images is widely used these days with the most practical form being digital documents instead of paper documents in the organizations and departments which make archives of digital documents for every individual. This means that with the rapid growth of new technologies, we are faced with the undeniable reality of increasing digital resources in organizations and departments. These resources are either digitized from paper documents or are in electronic format from the beginning.

In digital communications on internet, however, all data is visible and available to users. As a result, one of the most significant challenges is how to manage documents and how to develop strategies to protect these resources against any damages or alterations. The formation of Electronic Document Management Systems (henceforth EDMS) is in line with these concerns and has a growing trend.

In recent years, special attention has been paid to the information security and effective actions have been taken with respect to issues such as how to collect, monitor, publish and maintain digital resources in various organizations. Classification of documents and definition of users' access are common security actions in such systems. But whether these measures meet the needs of organizations is still debating. The growing number of attacks on text/image databases is one of the major concerns for the owners of electronic documents. These kinds of attacks usually have illegal, dangerous, and corrupt purposes. Therefore, storing and transferring document images must be secured. A few crucial questions might be arisen which is how the security of EDMS can be improved, and whether common methods of enhancing the security of digital images such as steganography and data encryption are solely able to address this important challenge. In this research, we suggest that to increase the security of digital images that are always exposed to various risks a combination of steganography, data mining and cryptography methods should be used. In this way, any changes in the image can be easily detected. We suggest using digital image processing techniques to extract unique and important information from images to form the desired text for steganography. Prior to embedding of the text inside the images, the text is encrypted by stream cipher which is one of the symmetric key encryption methods.

Some of the most important innovations of this research are:

1. Providing a framework based on two techniques: steganography, cryptography.
2. Using image processing to extract important features of images.
3. Generating a dynamic message/text based on a combination of two types of data, including: extracted features and additional information received from users.

4. Determining the optimal division point for separating the smooth and edge area in images.

5. Applying the suggested method in a real document management system in an organization.

6. Using different patterns for selecting adjacent pixels in the proposed algorithm.

Image processing is used to extract important information such as contrast, image size, number of black and white pixels, points correlation, among others from images. This information is then used in the generation of the text which is embedded in images. Before embedding stage, the text is encrypted by stream cipher. Finally, this method can be applied to a real document management system in an organization.

Some of the advantages of the proposed method are but not limited to the simplicity of the implementation and compatibility with the requirements of organizations on the one hand and an increase in the capacity of data embedding on the other hand.

The paper is organized as follows. Section 1 includes the introduction and description of the topic. Section 2 presents theoretical framework of research and technical properties of steganography and cryptography methods. Section 3 describes the literature review of recent related works. Section 4 is devoted to the research methodology. Section 5 presents experimental results. Finally, we wrap up the discussion and conclusion in section 6.

2. Theoretical Framework

In this section, definitions of related subjects and explanation of the commonly used techniques are discussed.

2.1. Information Management

Organizations often require a variety of information from their customers such as their identity documents. This information might come in different forms and to achieve better results, there is always a need to organize and manage information. Information management is an economical, efficient, and effective coordination in the processes of production, control, storage, retrieval, and dissemination of information from external and internal sources, which is aimed at improving the performance of organizations [7]. In other words, it is a method which uses technology to collect and process information for the purpose of efficient management [12].

Nowadays, with the advent of digital technologies it is possible to change, adjust, display, and present various information. However, there are some risks similar to the world of paper and printing. This requires organizations to consider different issues which might affect the correct use and the sustainability of resources. In addition, the management and use of digital resources have been growing in importance in many aspects, and it is undeniable that most of the activities and intellectual and knowledge assets of organizations are in digital form these days [6]. Online dictionary of library and information science defines digital archives as "A system designed for locating, storing and providing access to digital materials over the long term" [18].

There are two types of digital archive resources: (i) paper documents such as books, magazines, newspapers, documents, reports, images, and maps that can be converted to digital files, and (ii) digital documents such as emails or documents that are produced by computer systems. Although there are many differences in the approach of people in expressing the type of digital content, it should be noted that all the mentioned sources must be converted to digital format to be managed by archive management systems.

There are always many threats and risks to digital archives. Therefore, one of the most important issues in organizations is how to deal with risks and how to prevent threats. A list of potential threats in this area include [20]: (i) media, hardware and software failures, (ii) communication errors, (iii) network service failures, (iv) media, hardware and software obsolescence, (v) operators' errors, (vi) natural disasters, (vii) external and internal attacks, (viii) economic or organizational failures.

EDMS has a very important role in organizations and institutions nowadays. Part of the service delivery process is subject to the use of electronic documents. Electronic documents are significantly different from ordinary documents. A document is a collection of information about a subject that is represented by a variety of common symbols and with an understandable structure to humans [22]. According to a broader definition, an electronic document refers to any types of texts, sounds, images and graphics that are produced in the computer environment with tools such as fax, telex, e-mail, cameras, scanners, videos, keyboards and is considered as a piece of evidence [5]. A collection of electronic documents related to a person, or an object constitutes an electronic record.

As it is implied from the definitions above, EDMS can be introduced as a set of activities such as organizing, managing and controlling of electronic documents from production to operation. According to the definition provided by the International Council of Achieves of document storage system: Electronic Document Management is an information system that is designed for purposes such as storage and retrieval of documents and is organized to monitor processes such as production, storage and access to documents while maintaining their authenticity and citation capability [27].

EDMS usually consist of three main parts: document input, control and processing and sharing. Figure 1 shows an example of such a system presented by

Valkonen [26].

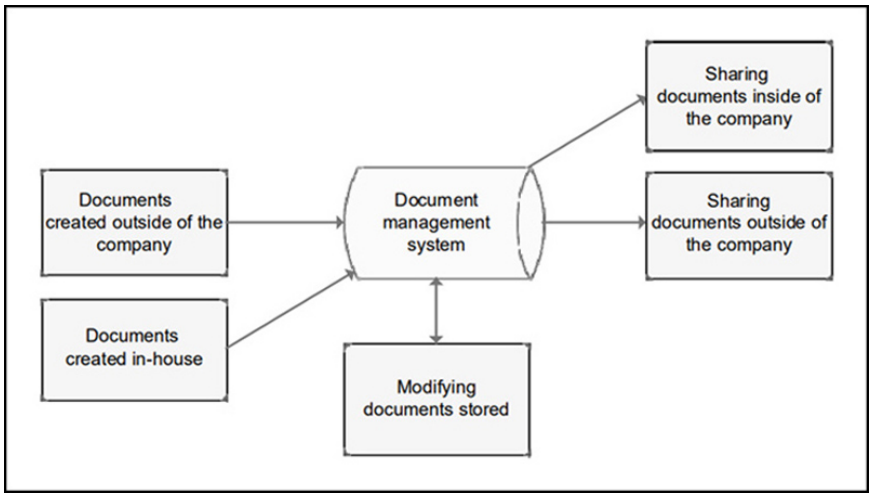

Figure 1: Document management system.

2.2. Information Security

A good document management system is a one that supports documents security from generating to publishing. Therefore, one of the most important challenges in these systems is the issue of documents security because of their exposure to threats such as forgery, distortion, and alteration. According to the security guidelines published by the National Institute of Standards and Technology (NIST), security is "the protection of an information system to achieve usable goals in order to maintain the integrity, accessibility, and confidentiality of information system resources [24]. In other words, security is a mechanism through which information and services are protected against unauthorized access, alteration, or destruction.

Due to the extensive use of digital resources in organizations, document management plays a new and vital role in promoting ecommerce-based services these days. Meanwhile, there are various ways to improve security in digital images the simplest of which is to categorize documents or define users' access and most importantly making use of steganography, watermarking and cryptography techniques which will be addressed in this study.

Effective implementation of document management systems plays an important role in creating a virtual work environment and changing the capabilities of the organization and its workforce. Obviously, information that passes through insecure channels is easily intercepted which is why the use of cryptography, watermarking and steganography methods plays an important role in information security [16].

2.2.1 Watermarking and Steganography

Watermarking is a way to embed specific information in digital media such as images, music and films [19]. Watermarking is a digital signal which carries the information that is used in various cases. The most important application of watermarking is in protecting the copyright of digital media and verifying the validity of their content. In general, there are two types of digital watermarking, visible and invisible watermarking.

Steganography with a history of more than thousand years, is one of the best ways to increase the security of electronic documents against attacks that lead to distortion or alteration. Concealing secret messages or military orders were widespread in ancient civilizations, and special methods have been devised throughout history based on existing technologies. In a simple definition, steganography is the art of hiding information in digital media such as images, sounds, and texts. In other words, steganography is a technique for transmitting confidential information in a way to avoid suspicion [3].

Steganography consists of three main elements: the media, the text or message, and the encryption key in which the key is optional. The process of embedding a message in the form of a mathematical equation is as follows:

$$S = Em(C, M, K),$$

wherein C is the cover image, M is the message and k is the key.

The mathematical equation of extracting a hidden message from the image (the stego image) is through:

$$M = Ex(S, K),$$

wherein S is the stego image and K is the key.

The text or message obtained from the extraction step should be the same as the message hidden in the stego image.

Watermarking is mostly used for copyright purposes and disclosure of unauthorized copies, whereas steganography covers a wider range of applications from sending hidden information via a medium to the audience to prove ownership, and most importantly the invisibility of information embedded in the media.

2.2.2 Cryptography

Cryptography is another tool to ensure information security. There are various algorithms for encrypting data which are generally divided into two main categories: symmetric key and asymmetric key. In symmetric key cryptography, decryption and encryption are performed using a single key, whereas in asymmetric key cryptography decryption and encryption operations are performed using different keys [23].

In this paper, the stream cipher method which is a symmetric key method is used. In this type of encryption, byte stream of the desired message is combined with an 8-bit keystream generated by pseudo-random generator using the XOR operator. The keystream is combined with the text stream one byte at a time using the XOR operation.

For example, if the next generated byte is 01101100 and the next text byte is 11001100, then the resulting cipher text byte is :

$$\begin{array}{rr} 11001100 & \text{(Text)} \\ +\,01101100 & \text{(key)} \\ ------ & (1) \\ 10100000 & \text{(ciphertext)} \end{array}$$

3. Literature Review

In the last decade, several studies have been conducted on how to use steganography, and cryptography methods to increase information security. The following paragraphs discuss some of these studies briefly.

In the method presented by Bhuiyan et al. [9], the calculation of the difference between values of adjacent horizontal, vertical, and diagonal pixels in $2 * 2$ blocks and the use of a fixed table specifying the upper and lower ranges has been suggested for steganography. In this method, the desired message is encrypted and normalized prior to steganography. The results show an increase in the security as well as a slight decrease in images quality after steganography. Data embedding capacity is low in this method.

Shaik and Amitharajan [21] present a combination of steganography, cryptography and process randomization methods to increase the security of digital images and evaluate the proposed method using Peak Signal to Noise Ratio (henceforth PSNR) criteria and robustness. In the hybrid method provided by Astuti et al. [4], the text before steganography in the least significant bit (LSB) is encrypted three times with the XOR operator and three most significant bits (MSB) are used as the keys. The results show the desired value of PSNR with no decrease in the image quality. One of the strengths of this method is the simplicity of algorithm. In another study by Abdur Razzaq et al. [2], a combination of steganography and watermarking techniques is used to increase the image security. In this research, XOR operator is used for encryption and LSB method is used for steganography. This method is evaluated by different criteria and the results show the effectiveness of the suggested method, while the reduction in the image quality is very low considering the PSNR value. The solution provided by Miñon et al. [17] is more useful for managing and supervising university documents on an Intranet platform with following sections: login module, document management, monitoring, notification, settings, and messaging. The results of this study show the improvement

of security weaknesses.

Zhaotong and Ying [15] present a hybrid method based on the difference of adjacent pixel values in $2 * 2$ blocks, in which an evolutionary and metaheuristic Particle Swarm Optimization algorithm is used. This algorithm is inspired by the behavior of fish and birds and aims to find the optimal value through an iterative process based on random solutions. Bandyopadhyay and Bhattacharjee [8] use a combination of three techniques (i.e. spatial domain image modification, discrete cosine transforms and image scrambled using modified Arnold Transform) to generate the text of the message. The results of this study are evaluated with PSNR criteria. Delenda and Noui [11] intruduce a new steganography algorithm based on a linear algebraic tool, the polar decomposition, for hiding secret data in an image. The results of this method show the appropriate concealment capacity and high resistance and security against common types of attacks. Hussein et al. [14], suggest a new solution for data embedding in images By combining pixels value differencing and rightmost digit replacement methods. In this method, the cover image is partitioned as non-overlapping two-pixel blocks and the difference between two adjacent pixels is calculated. The data embedding method and number of hidden bits is different depending on whether the difference value is related to the lower or upper area of the defined table. The method suggested by Soria-Lorente and Berres [16] is based on the compression and entropy threshold techniques. In the text encryption section, they use a public key and a private key to generate a pseudo-random binary string. The results show the improvement of security and the desired values of evaluation criteria. One of disadvantages of this research is that the data embedding capacity is not investigated. In the method presented by Chang et al. [10], the calculation of the difference between values of adjacent pixels in $2 * 2$ blocks is suggested. This method is evaluated by different criteria and the results show the effectiveness of the suggested method.

4. Research Methodology

In this paper, appropriate features are extracted from digital documents using image processing techniques. The text, which is a combination of mentioned features, is encrypted by the stream cipher method before being hidden in the image. This process is checked using valid and common tests. The type of research is a survey with a modeling approach and is considered applied according to the purpose of the research. The research population is the digital documents of insureds of the Social Security Organization, so the method of collecting information in this research is the use of images from the mentioned databases. The sample size is 105 electronic records from digital documents of Tehran branches. The methodology of this research is done in five stages as can be seen in Table 1.

Before describing the research stages, it is necessary to provide a brief description of the research community, which is the digital documentation of the insureds of the Social Security Organization. In the Social Security Organization, which

Table 1: Steps of research methodology.

Name	Target	Necessary actions
Feature extraction	Extracting the unique features of images	Converting images to black & white and extracting the required features
Desired text creation	Creating composite text	Combining: 1.extracted features 2.additional information
Stream cipher	Preventing the discovery of hidden data in the image	Encrypting the desired text with the stream cipher method
Steganography	Hiding the text in the cover image	Embedding the text in the image with the proposed method
The proposed method evaluation	Evaluating the results	Calculating MSE, PSNR and capacity for steganography evaluation

is more than 60 years old, the digital archive project started in 2017 and most of its documents have become digital in recent years. To register or receive services, the clients of this organization must provide documents such as identity cards, national cards, military service cards, marriage certificates and medical certificates to name a few. These documents are kept in the form of an electronic record after being scanned onto the system. Therefore, each client has an electronic record. A sample of digital documents in this organization is shown in Figure 2.

We can now examine how to perform the steps mentioned earlier based on the conceptual model shown in Figure 3.
Each of the steps mentioned in Table 1 includes a set of processes that are used to implement the suggested method. The results obtained in each step are used in completing the operational steps and analyzing the proposed method. In the following section, the results of each step are presented along with necessary explanations.

4.1 Feature Extraction

Each digital image is composed of a set of general information items such as size, type, location, date of production, and in a closer look it has special and unique information features that can be obtained only with image processing techniques. In this research, 12 features for each image that include number of pixels, contrast, number of corners, resolution, number of key points, homogeneity, correlation of points, energy, width, height, are extracted.

Contrast measures the intensity variation between the reference pixel and adjacent pixel. Correlation measures, how the reference pixel is related to its neighbor pixel. Energy defines the measure of sum of squared elements. When pixels are

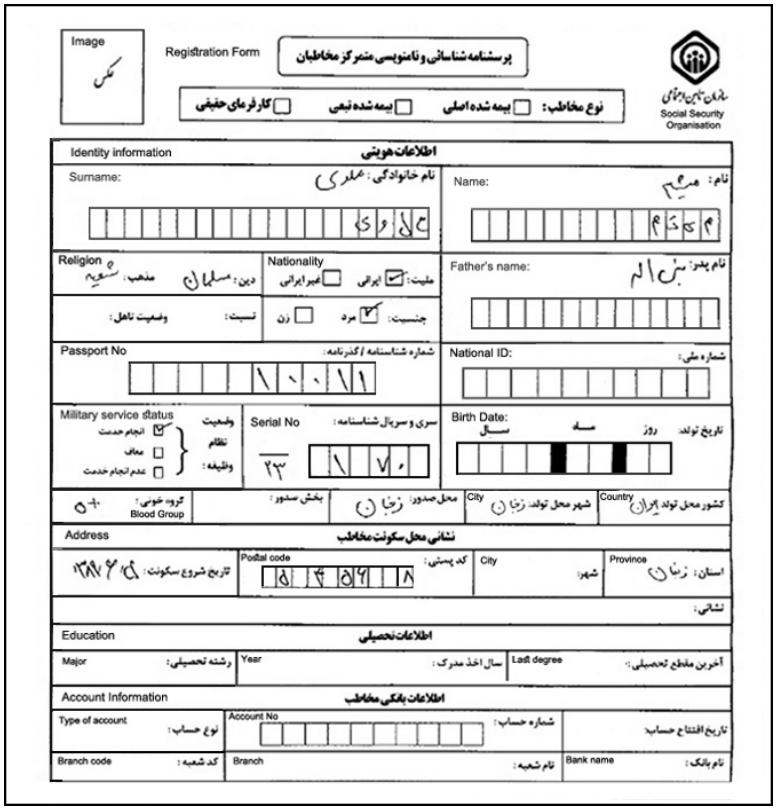

Figure 2: Sample of a digital documents of social security organization.

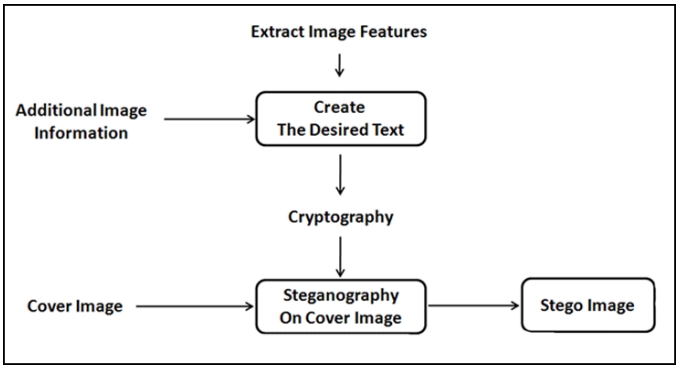

Figure 3: Conceptual model of the proposed framework.

very similar, the energy value will be large [25].

Homogeneity measures how close the distribution of a pixel is in a gray level co-occurrence matrix. This value will be inversely proportional to contrast [13].

For example, in Table 2 shows some of the features extracted from the image of a national card.

Table 2: Features extracted from national card.

Width	Height	Pixels	Contrast	Corners	Key Points
388	605	234740	0.5369	1231	682
Homogeneity	Correlation	White pixels	Black pixels	Resolution	Energy
0.9373	0.8669	13236	220512	96	0.7215

These unique features are used to generate the dynamic text for steganography.

4.2 Creating Text

In the electronic record of each insured, there are various documents. One of these documents, which is called a checklist, contains information such as: name and insurance number, number of documents and date of record formation. In this paper, this document will be used as a cover image. To create the desired message for steganography, information such as insurance number, insured name, number of documents, date, social security branch code, with unique features obtained from each image are combined.

For instance, the text that is generated for an insured with the insurance number of 0015024514 and features specified in Table 3 is as follows:

Table 3: Features of 3 documents of an insured.

Document number	Date	Branch code	Pixels	Corners	Contrast	White pixels	Black pixels	Correlation
1	13970204	25	410495	1946	2, 11	16182	307806	0.3972
2	13970204	25	442339	4844	2, 83	62154	378846	0.7256
3	13970317	13	322525	4324	3, 74	55463	265913	0.7051

By combining the insurance number and the number of documents with the features of documents, the desired text to be embedded is created:
$015024514 - 3 - 1 - 13970204 - 25 - 410495 - 1946 - 2.11 - 16182 - 307806 - 0.3972 - 2 - 13970204 - 25 - 442339 - 4844 - 2.83 - 62154 - 378846 - 0.7256 - 3 - 13970317 - 13 - 322525 - 4324 - 3.74 - 55463 - 265913 - 0.7051$

In the above example, only some of the features have used for simplicity.

4.3 Steganography

In this section, a new method is presented for the steganography of the encrypted text which is based on calculating the difference between the value of the gray-level of two adjacent pixels in 4×4 blocks. A division point from 0 to 255 must be considered to separate the smooth area and the edge of the gray surface. This division point can be selected from numbers such as 1,3,7,15,31,63 or 127. These points are calculated by equation given in (4).

$$2^n - 1, \qquad for \quad n = 1, 2, 3, 4, 5, 6, 7. \tag{4}$$

Naturally any division point that is selected affects the results of the algorithm which is why the optimal division point should be selected based on the purpose of the research. In the proposed algorithm, number 7 is selected for the division point.

The proposed embedding algorithm is as follows:

Input: Cover image C, secret information Msg, division point Div.
Output: Stego image S.

Step 1: If the dimensions of C is $M \times N$ with a separate arrangement of 4-pixel rows, the image is shown as $L = (M \times N)/4$ so that each cover image is $4 \times L$.
Step 2: Divide the $4 \times L$ cover image into non-overlapping 4×4 blocks.
Step 3: The horizontal and vertical edges of each 4×4 block is obtained by calculating the difference values of two adjacent pixels (horizontal or vertical) based on the pattern shown in Figure 4 and through following calculations.
For each pair of horizontal adjacent pixels:

$$d_i = |p_i - p_{i+1}|, \qquad i = 1, 3, 13, 15.$$

For each pair of vertical adjacent pixels:

$$d_i = |p_i - p_{i+4}|, \qquad i = 5, 6, 7, 8.$$

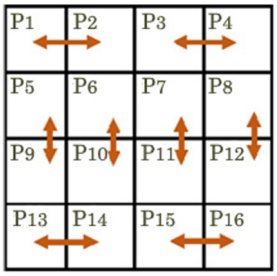

Figure 4: Pattern of calculating in the adjacent pixels.

Step 4: Find the optimal value of region R_i from d_i in which $R_i = \min(u_i - k)$ using the following conditions:
$u_i > k$, $k = |d_i|$, $R_i \in [l_i, u_i]$ and optimal region R_i for all value of $1 \leqslant i \leqslant n$.
Step 5: Compute and read number of the secret bits as $t = |\log_2{}^{w_i}|$ from Msg and convert into decimal value denoted as b. ($w_i = U_i - L_i + 1$)
Step 6: Calculate the new difference value d which is given by $d_{new} = L_i + b$.
Step 7: If R_i is not in the range of 8 to 255 bits go to step 9.
Step 8: Modify the P_i,P_{i+1} by following equation:

$$(P'_i, P'_{i+1}) = \begin{cases} (P_i) + \left\lceil \frac{m}{2} \right\rceil, (P_{i+1}) - \left\lfloor \frac{m}{2} \right\rfloor, \ if\ P_i > P_{i+1}, d_{new} > d_i \\ (P_i) - \left\lfloor \frac{m}{2} \right\rfloor, (P_{i+1}) + \left\lceil \frac{m}{2} \right\rceil, \ if\ P_i > P_{i+1}, d_{new} > d_i \\ (P_i) - \left\lceil \frac{m}{2} \right\rceil, (P_{i+1}) + \left\lfloor \frac{m}{2} \right\rfloor, \ if\ P_i \geqslant P_{i+1}, d_{new} \leq d_i \\ (P_i) + \left\lceil \frac{m}{2} \right\rceil, (P_{i+1}) - \left\lfloor \frac{m}{2} \right\rfloor, \ if\ P_i < P_{i+1}, d_{new} \leq d_i \end{cases}$$

Step 9: 6 bits of the encrypted text message stream is read and the message is embedded using the 3-LSB method. (convert P_i,P_{i+1} to be P'_i,P'_{i+1} by substituting 3-LSB of P_i,P_{i+1}). If $|p'_i - p'_{i+1}| > Div$ then readjust P'_i,P'_{i+1} by following equation:

$$(P'_i, P'_{i+1}) = \begin{cases} P_i - 8, P_{i+1} + 8, \ if\ P_i \geqslant P_{i+1}, \\ P_i + 8, P_{i+1} - 8, \ if\ P_i < P_{i+1}. \end{cases}$$

Step 10: Repeat steps 3 to 9 until the end of the message stream.

5. Experimental Results

In this section, we evaluate the performance of the proposed steganography scheme by running several experiments. The visual quality and embedding capacity are considered performance measures. The proposed method is tested using MATLAB 2018 on standard images of Lena, Baboon and Peppers (Figure 5) with a size of 512Œ512 pixels, which are commonly used in previous studies.

Figure 5: Images of Lena, Baboon and Peppers.

5.1 Analysis of Stego Images Visual Quality

Data embedding in a cover image usually alters it, and there may be some changes in pixel values that affect the visual quality of the stego image. These changes

must be investigated since they directly affect the imperceptibility of the final appearance of the stego image.

To investigate information steganography, two metrics of the Peak Signal to Noise Ratio (PSNR) and Mean Squares of Errors (MSE) are used on MATLAB 2018.

The MSE index is the mean squared error between the cover image and the stego image which is calculated through the Equation (2).

$$MSE = \frac{1}{M * N} \sum_{i=1}^{M} \sum_{j=1}^{N} (c_{i,j} - s_{i,j})^2, \tag{2}$$

where M and N are the number of horizontal and vertical pixels respectively in the cover image and $c_{i,j}$, $s_{i,j}$ are the values of pixels in the cover and stego image respectively. MSE is the most common estimator of the image quality measurement metric. It is a full-reference metric with values closer to zero being better.

The PSNR index is used to measure the quality of the stego image, and its value is obtained through the equation given in (3). A higher PSNR value provides a better image quality.

$$PSNR = 10 * \log_2 \frac{\max^2}{MSE}, \tag{3}$$

where max is the maximum number of possible pixel values of the image. When pixels are represented using 8 bits per sample, max is 255.

Table 4 shows the comparison of the results of proposed method with some other methods. The PSNR value in three images are greater than 30, indicating

Table 4: Comparison of the results of the proposed method with other methods.

Cover image	Hussain et al		Chang et al		Bhuiyan et al		Proposed method	
	PSNR (db)	Capacity (Bytes)	PSNR (db)	Capacity (Bytes)	PSNR (db)	Capacity (Bytes)	PSNR (db)	Capacity (Bytes)
Lena	39.09	563152	38.89	75836	38.91	74047	38.94	518624
Baboon	35.06	631783	33.63	82407	39.30	74047	35.99	528008
Peppers	39.30	562573	38.50	75579	39.15	74047	37.55	524532

that the distortion of stego images is undetectable to human vision [1]. The results of the capacity in three images show the acceptable performance of the data embedding process.

May be seem that the results of proposed method is not much different from other methods, but it should be noted that the proposed method is based on the optimal results for capacity and PSNR. If the purpose of this research is simply to obtain a PSNR value higher than other methods, This will be achieved by changing the division point.

For example, by selecting the division point of 3 for Lena image, the PSNR value exceeds 40.

Another common method for evaluating embedding data is to compare the original images histogram with the stego image. Figure 6 compares the Lena images before and after embedding.

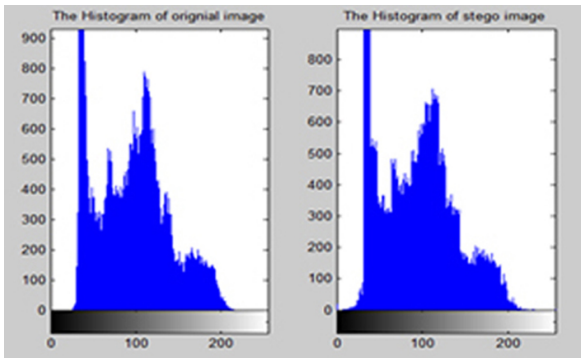

Figure 6: Comparison of Lena image histograms before and after steganography.

As can be seen, there are some changes in some of the columns in the right image. But in general, these two histograms are very similar to each other. Which means that the image quality has not decreased much after embedding the data.

5.2 Impact of Changing the Division Point on PSNR and Capacity

In this section, the effect of changing the division point on the embedding capacity and PSNR is investigated. One of the most important parts of the proposed algorithm is selecting the dividing point to separate the smooth area and the edge of the gray surface.

This division point can be selected from numbers such as $1, 3, 7, 15, 31, 63$ or 127 and any division point that is selected affects the results of capacity and PSNR.

To further investigate this issue, the suggested algorithm is tested on Lena image with different division points. The results of this experiment are given in Table 5.

As can be seen, for smaller division points a higher PSNR is obtained whereas for larger division points the message embedding capacity increases. If the aim of the study is only to obtain the desired image quality, the optimal point that has the highest PSNR should be selected after the implementation of the proposed method. However, if we want to consider both the image quality and the capacity of embedding the message at the same time the $Div = 7$ is the most appropriate case.

Table 5: The results of running the proposed algorithm with different division points on the Lena image.

Divison Point	MSE	PSNR	Capacity
1	5.36	40.84	227292
3	6.19	40.22	330360
7	8.29	38.94	518624
15	14.75	36.44	758320
31	32.65	44.46	1025576
63	173.29	25.74	1300680
127	655.35	19.96	1571488

5.3 Review of the Proposed Method on EDMS of Social Security Organization

To test the applicability of our method in an electronic document management system, the proposed algorithm is tested on the digital documents of clients of the Social Security Organization and the data analysis is reviewed based on the above criteria. The results indicate that the quality of images does not decrease after steganography and an appropriate capacity is retained to embed information in images.

Table 6 shows the results of the implementation of the mentioned method on three types of common documents available in the digital files of an insured of the Social Security Organization.

Table 6: Results of the proposed method on the documents of an insured.

Cover Image	MSE	PSNR	Capacity
Checklist	15.56	36.21	537084
National cards	17.18	35.78	553584
Identity cards	40.51	32.06	568804

6. Discussion and Conclusion

Ensuring the security of digital information is a major challenge for organizations. In this paper, a framework was presented for the security of digital images. It is a fusion of two security methods (i.e. steganography, and cryptography). This framework mainly includes three phases. In the first step, extracting unique features of each image generate a dynamic message. In the next step message

encryption reduces the possibility of detecting the generated text. Finally, the process of embedding encrypted message in cover images through using the new presented algorithm can increase the risk resistance.

The experimental results reveal that the visual quality of cover images and the capacity of data embedding in the proposed method is acceptable and the evaluations performed with different criteria confirm this issue.

Another important point to note is that, any small changes in a digital document even up to one pixel can be detected in the proposed method due to the inclusion of various statistical information in the images. Therefore, using this method in document management systems can play an effective role in resolving security concerns.

In summary, the role and significance of all components of the suggested framework which is developed to create more security for the images can be emphasized.

The combination of the mentioned techniques and the dynamics of the information extracted from images can be effective in increasing the security of digital images in electronic document management systems.

Among the limitations of this research are lack of similar research in this field especially in its practical form, limitations of access to digital images, complexity of each process and their combination with each other.

Conflicts of Interest. The authors declare that there are no conflicts of interest regarding the publication of this article.

References

[1] M. Abdel hameed, M. Hasaballah, S. Aly and A. I. Awad, An adaptive image steganography method based on histogram of oriented gradient and PVD-LSB techniques, *J. IEEE Access* **7** (2019) 185189 – 185204.

[2] M. Abdur Razzaq, M. A. Baig, R. A. Shaikh and A. A. Memon, Digital image security: fusion of encryption, steganography and watermarking, *J. Adv. Comput. Sci. Appl.* **8** (5) (2017) 224 – 228.

[3] A. S. Ansari, M. S. Mohammadi and M. T. Parvez, A comparative study of recent steganography techniques for multiple image formats, *IJCNIS* **1** (2019) 11 – 25.

[4] Y. P. Astuti, D. R. I. M. Setiadi, E. H. Rachmawanto and C. A. Sari, Simple and secure image steganography using LSB and triple XOR operation on MSB, In: *Int. Conf. Info. Commun. Tech.* (ICOIACT), Yogyakarta, Indonesia (2018) 191 – 195.

[5] G. Azizi (in Persian), *Electronic Document Management*, Organization of Documents and National Library, I. R. Iran, 2011.

[6] N. Beagrie, Digital curation for science, digital libraries, and individuals, *Int. J. Digit. Curation* **1** (1) (2006) 3 – 16.

[7] D. P. Best, The future of information management, *Records Manag. J.* (2010) 61 – 71.

[8] H. Bhattacharjee and S. K. Bandyopadhyay, Frequency domain approach of image steganography, *Int. J. Innovative Res. Info. Sec.* **3** (2) (2016) 9 – 16.

[9] S. S. N. Bhuiyan, N. Abdul Malek, O. O. Khalifa and F. D. Abdul Rahman, An improved image steganography algorithm based on PVD, *Indones. J. Electr. Eng. Comput. Sci.* **10** (2) (2018) 569 – 577.

[10] K. Chang, C. Chang, P. S. Huang and T. Tu, A novel image steganographic method using Tri-way pixel-value differencing, *J. Multimedia* **3** (2) (2008) 37 – 44.

[11] S. Delenda and L. Noui, A new steganography algorithm using polar decomposition, *Inf. Secur. J.: A Global Perspective* **27** (3) (2018) 133 – 134.

[12] Gartner inc, [Online]. Available: https://www.gartner.com/en/information-technology/glossary/im-information-management, Accessed 20 jan (2020).

[13] T. Haryanto, A. Pratama, H. Suhartanto, A. Murni, K. Kusmardi and J. Pidanic, Multipatch-GLCM for texture feature extraction on classification of the colon histopathology images using deep neural network with GPU acceleration, *J. Comput. Sci.* **16** (3) (2020) 280 – 294.

[14] M. Hussain, A. W. AbdulWahab, A. T. S. Ho, N. Javed and J. Ki-Hyun, A data hiding scheme using parity-bit pixel value differencing and improved rightmost digit replacement, *Signal Processing: Image Commun.* **50** (2017) 44 – 57.

[15] Z. Li and Y. He, Steganography with pixel value differencing and modulus function based on PSO, *J. Info. Secur. Appl.* **43** (2018) 48 – 49.

[16] A. S. Lorente and S. Berres, A secure steganographic algorithm based on frequency domain for the transmission of hidden information, *Secur. Commun. Networks* **2017** (2017) 5397082.

[17] J. D. F. Miñon, C. M. A. Lim, J. A. L. Morano, R. F. Fajutagana and B. S. Fabito, An intranet-based document management and monitoring system framework: A case for the National University Quality Management Office, 2016 *IEEE Region* 10 *Conf.* (TENCON) (2016) 2262 – 2267.

[18] ODLIS, [Online]. Available: https://www.abc-clio.com/ODLIS/odlis-d.aspx, Accessed 20 jan (2020).

[19] L. Robert and T. Shanmugapriya, A study on digital watermarking techniques *Int. J. Recent Trends Eng.* **1** (2) (2009) 223 – 225.

[20] D. S. H. Rosenthal, T. Robertson, T. Lipkis and V. Reich, Requirements for digital preservation systems: a bottom-up approach, *D-Lib Magazine* **11** (11) (2005). Available: http://www.dlib.org/dlib/november05/rosenthal/11rosenthal.html

[21] A. Shaik and R. Amitharajan, Data security through data hiding in images: a review, *J. Artif. Intell.* **10** (1) (2017) 1 – 21.

[22] R. H. Sprague, Electronic document management: challenges and opportunities for information systems managers, *MIS Quarterly* **19** (1) (1995) 29 – 49.

[23] W. Stallings, *Cryptography and Network Security Principles and Practice*, 5th ed., Prentice Hall, New York, 2011.

[24] G. Singh, Supria: A study of encryption algorithms (RSA, DES, 3DES and AES) for information security, *Int. J. Comput. Appl.* **67** (19) (2013) 33 – 38.

[25] B. Thamaraichelvi and G. Yamuna, Gray Level Co-Occurrence Matrix features based classification of tumor in medical images, *ARPN J. Eng. Appl. Sci.* **11** (19) (2016) 11403 – 11414.

[26] J. Valkonen, *Document Management for SmallMusiness*, Master's Thesis, Turku University of Applied Sciences, 2015.

[27] B. Zarrinkelki (in Persian), Government organisations and electronic documents, *Treasures of Documents* **20** (2) (78) (2011) 100 – 120.

Bashir Omrani Harzand
Department of Industrial Management,
Central Tehran Branch,
Islamic Azad University,
Tehran, I. R. Iran
e-mail: omrani@hotmail.com

Mohammadreza Motadel
Department of Industrial Management,
Central Tehran Branch,
Islamic Azad University,
Tehran, I. R. Iran
e-mail: moh.motadel@iauctb.ac.ir

Ali Broumandnia
Department of Computer Engineering,
South Tehran Branch,

Islamic Azad University,
Tehran, I. R. Iran
e-mail: broumandnia@gmail.com