Proceedings of the Thirtieth International Florida Artificial Intelligence Research Society Conference

A Robust Watermarking System Based on Formal Concept Analysis and Texture Analysis

Musab Ghadi,[1] Lamri Laouamer,[1,2] Laurent Nana,[1] Anca Pascu,[1]
[1] Lab-STICC (UMR CNRS 6285), University of Brest
CS 93837, 29238 Brest Cedex, France
[2] Department of Management Information Systems, Qassim University
P.O. Box 6633, Buraidah, 51452, KSA
Emails: e21409716@etudiant.univ-brest.fr, laoamr@qu.edu.sa, {Laurent.Nana, Anca.Pascu}@univ-brest.fr

Abstract

Designing new watermarking models for the authentication of transmitted images over wireless multimedia sensor networks is an essential need. Indeed, wireless networks require considering several constraints including execution time, robustness, and fault tolerance against different attacks. Computational intelligence and knowledge discovery approaches, such as rough set and Formal Concept Analysis (FCA), make it possible to efficiently exploit the host image characteristics to identify interesting locations for watermark embedding. In this paper a robust watermarking approach based on FCA is proposed. It derives relationships between the image's blocks described through a set of texture features, and uses these relationships to identify the more interesting blocks for watermark embedding. The experiment's result proves the efficiency of the proposed technique in terms of imperceptibility and robustness against different attacks.

Introduction

The emergence of many applications that transmit digital images over wireless multimedia sensor networks requires maintaining the security and authentication. The environmental monitoring, remote sensing images, and medical image analysis (Sudhir 2015) are widely used applications based on transmitting images across wireless networks securely. The characteristics of targeted image can be used to infer a meaningful knowledge from the explicit image information (Sawant 2010). The color representations, the texture/smooth nature, and the structure of image's surface/background are important characteristics of the digital images. The Formal Concept Analysis (FCA) is a fundamental technique in the field of knowledge discovery and image mining (Priss 2006). It has a significant role in different applications including e-healthcare system (AlShaikh et al. 2016), software engineering (Tilley 2005) and linguistic (Priss 2005). Few applications in the watermarking field were recently proposed (AlShaikh et al. 2016; Quist-Aphetsi et al. 2015). In this paper, the FCA theory is used to mine the texture features of targeted images. The DC, skewness, kurtosis, and entropy are four descriptive texture features that may be analyzed by FCA to find a meaningful knowledge, which helps in determining highly textured blocks of targeted image. These blocks are preferable for a robust embedding of the watermark, since human's eye is less sensitive to modifications in textured regions (imperceptibility). Four metrics are considered to evaluate the performance of the proposed model with respect to the constraints of wireless networks: the perceptual quality of embedded image through Peak Signal-to-Noise Ratio (PSNR) and Structural SIMilarity (SSIM), the robustness of embedded watermark through Bit Error Rate (BER) and Correlation Coefficient (CC). The rest of this paper is organized as follows: the related work is presented in section 2. Section 3 presents a background of FCA and texture analysis of digital images. Section 4 presents the system model and section 5 deals with the experiments result. The paper ends by a conclusion in section 6.

Related Work

This section explores some of the suggested watermarking models based on image mining, computational intelligence and knowledge discovery techniques, in order to provide image authentication. Most of them used the image texture analysis features as main input to the different techniques of mining and knowledge discovery.

The authors in (Quist-Aphetsi et al. 2015) proposed a digital image watermarking approach in the spatial domain based on image features built from FCA. The first step of the watermarking approach is the extraction of image features using FCA. For the application of the FCA process, the arithmetic mean, the standard deviation, and the entropy of the image are used. The concepts resulting from the application of the FCA process are used to generate the watermark. A change in a pixel value results in a change in the concepts generated from the image using the FCA. The proposed watermarking approach was combined with a ciphering approach based on pixel displacements. The concepts generated using FCA are maintained throughout the process. This makes it possible to authenticate images from their ciphered version, using the FCA concepts.

In (AlShaikh et al. 2016) the authors proposed a watermarking approach in the spatial domain based on FCA, in order to improve the robustness of medical images. After isolating the zeros at the edge of the original image, four

12×12 sub-matrices are first extracted at the four corners of the rest of the image. Each sub-matrix is then converted into a matrix of size 144×8 resulting from the conversion of each pixel value of the sub-matrix into its 8-bits representation. The formal concepts are then built from the binary representation of the sub-matrices, the extents been row numbers and the intent been column numbers. The watermark is built from information extracted from the header of the DICOM (Digital medical imaging and communication in medicine) image. It is inserted at the locations obtained from the application of the FCA process previously described. The experimental results show the extraction of the embedded watermark image with high quality, low payload, and low computational complexity from the attacked watermarked image after some attack scenarios.

Recently, the authors in (Ghadi et al. 2016) proposed a semi-blind associative watermarking model to achieve the image authentication based on texture analysis and frequent patterns mining. The proposed model mines four textured features (DC coefficient, skewness, kurtosis, and entropy) to extract most frequent pattern. The extracted pattern is used to identify more textured blocks for embedding watermark. The experiments result proved the efficiency of the model in terms of BER, CC, and PSNR against different attacks. The advantage of this approach comes from the fact that embedding a watermark in highly textured regions would be less sensitive to the attacker's eye (preserve the perceptual quality of host image) and then will maintain good robustness against different attacks.

Background

Principle of FCA

FCA is a technique used to investigate and analyze image characteristics, in order to find meaningful and comprehensive knowledge (Poelmans et al. 2010). It was developed in the field of data mining, knowledge representation, and knowledge discovery in databases (Alqadah and Bhatnagar 2011). FCA manipulates a data matrix, which combines set of objects and set of attributes, to find the set of all objects that share a common subset of attributes and the set of all attributes that are shared by one of the objects.

FCA theory relies on different notions. The basic notion in FCA is a formal context defined as a triple β=(G,M,I), where *G* is a set of formal objects, *M* is a set of formal attributes, and *I* is a binary relation called incidence such as $I \subseteq G\times M$. The notation *gIm* stands for $(g,m)\in I$, which is read as: the object *g* has the attribute *m* (Poelmans et al. 2010). A pair (X,Y) is a formal concept (FC) of (G,M,I) if and only if: $X \subseteq G$ (*X* is a subset of objects of *G*), $Y \subseteq M$ (*Y* is a subset of attributes of *M*), X'=Y (X' is the set of attributes in *M* such that all objects in *X* have all attributes in X'), and X=Y' (Y' is the set of objects in G such that all attributes in *Y* fall under all objects in Y'). *X* and *Y* are respectively called the Extent and the Intent of the FC.

Texture analysis of digital images

In the literature, there are four descriptive features used commonly to define the texture property of image. These features include the DC coefficient, skewness, kurtosis, and entropy. All of these features are presented in (Ghadi et al. 2016) along with the algorithm used to discriminate between textured and smooth blocks.

The DC feature expresses the average information of the overall magnitude of the image, and it is used as a fine property to define the energy of a given image. High-energy image is more textured than low-energy image (Ghadi et al. 2016).

The skewness feature measures the degree of asymmetry distribution of gray-level intensities around the mean and is used to indicate the texture/smooth nature of image. The image is textured if it is dense towards the white (i.e. in the negative skewness) or it is dense towards the black (i.e. in the positive skewness), while the normal distribution does not express any knowledge regarding the texture or smooth nature (Das 2015).

The kurtosis feature measures the flatness of gray-level intensities around the mean and expresses the texture/smooth nature of the image. If the image has low gray-level intensities (i.e. it is dense toward the black), then the distribution of the gray-level intensities is flat around the mean, this means that the image has low kurtosis value and it is textured. While, if the image has high gray-level intensities (i.e. it is dense toward the white), then the distribution of gray-level intensities is peaky around the mean, this means that the image has high kurtosis value and the information content would be significantly low (i.e. smooth image) (Das 2015).

The entropy feature measures the uniformity of the distribution of gray-level intensities of the overall image, and indicates the magnitude of image's information. High entropy value means that the gray-level intensities are distributed randomly along the image and the image combines dispersant pixels' values, this indicates that the image has much information and it's close to be textured. While, low entropy value means that the distribution of gray-level intensities is uniform along the image and the image combines similar pixels' values, this indicates that the image has not much information and it's close to be smooth (Umbaugh 2011).

System Model

The proposed model suggests a robust watermarking scheme based on FCA. The FCA is used to deduce the texture features of targeted image (based on DC, skewness, kurtosis and entropy), and then to discover a meaningful knowledge that helps to identify highly textured blocks for embedding the watermark. The idea of embedding watermark in highly textured regions is correlated with HVS principles, where the attacker's eyes become less sensitive to any change in highly textured regions rather than smooth regions (Yang and Yin 2013). This in fact may lead to preserve perceptual image quality and to achieve high robustness.The pseudocode of the proposed model is presented in algorithm 1.

According to algorithm 1, the proposed model operates mainly through six steps that are illustrated in the following subsections.

Algorithm 1 The pseudo-code of the proposed model

Preliminary: defining the set x=$\{x_1, x_2, ..., x_n\}$ as texture features
Input: host image *I* sized M×N and watermark image *w* sized L×L
building the Transactions matrix and Boolean matrix
applying FCA to extract the set of formal concepts
computing the frequency of each object in formal concepts, as well as computing the mean and the median of all frequencies to assign maximum one as a threshold *T*
identifying a set of highly textured blocks based on *T*
embedding watermark (I,w)
extracting watermark (I_{wa},w)

Building the transactions and Boolean matrices

In this step, the targeted image is partitioning into L×L non-overlapping blocks and the values of the texture features for every block are computed to build the transactions matrix. Subsequently, the transactions matrix is transformed into a Boolean matrix based on the thresholds (DC_{mean}, $skewness_{mean}$, $Kurtosis_{T1}$, $Kurtosis_{T2}$, and $entropy_{mean}$). Figure 1 presents the structure of this step.

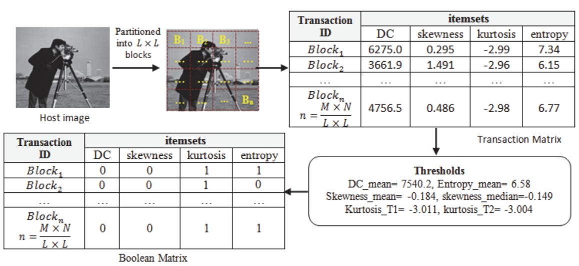

Figure 1: The structure of transactions and Boolean matrices.

Applying FCA to extract the formal concepts

In this step, FCA processes the Boolean matrix to extract the set of formal concepts. The resulting formal concepts present the relationships between the objects (blocks) and the attributes (texture features). Figure 2 shows the structure of formal concepts for a given Boolean matrix in (a), which consists of five objects and four attributes. (b) presents the 11 resulting concepts and (c) presents one of the 11 formal concepts that combines the set of objects {3, 5} as Extent and the set of attributes {DC, Entropy} as Intent.

Computing the frequency of each object in the formal concepts

The frequency of each object through all formal concepts is computed, then the mean and the median values of these frequencies are computed. The maximum between the mean and median defines the threshold *T*.

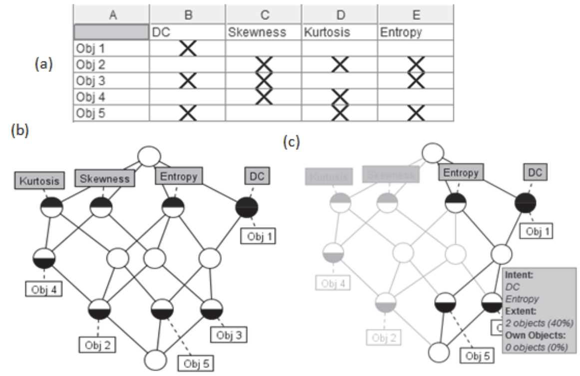

Figure 2: The structure of formal concepts.

Identifying a set of highly textured blocks based on threshold T

Any object (block) whose frequency is greater than the threshold *T* is considered highly textured. An object whose frequency is high in all formal concepts has a high ratio of attributes falling under it.

Watermark Embedding Process

All blocks identified as highly textured are considered in the embedding process. The embedding process is achieved by applying the linear interpolation technique. This technique gives the ability to maintain the invisibility of watermark in the targeted image by using a proper interpolation factor *t*. The range of interpolation factor is]0-1[, where using high interpolation factor *t* in the embedding equation ensures good level of invisibility. Algorithm (2) illustrates the watermark embedding process.

Algorithm 2 The pseudo-code of embedding watermark

Input: host image *I*, watermark image *w*, and interpolation factor *t*
for each $block_i \subseteq$ the set of highly textured blocks, $block_i \in$ I: block size=watermark size **do**
    $block_{iw}$ = (1/t)w+t×$block_i$: t=0.98   (1)
end loop
Output: watermarked image (I_w)

Watermark extraction process

The resulting watermarked image I_w, which holds the watermark data, is subject to channel errors and attacks due to the transmission across a public network. The extraction process in the receiver side is achieved to verify the authenticity of transmitted images. Algorithm (3) below, presents the watermark extraction process using the inverse form of linear interpolation technique. It uses the same interpolation factor as that used in the embedding process.

Experiments Result

This section presents the experiments result of the proposed system against different attacks. The experiments were con-

Algorithm 3 The pseudo-code of extraction watermark

Input: attacked watermarked image I_{wa}, the watermark image w, and interpolation factor t
for each block$_i$ $\subseteq$ the set of highly textured blocks, block$_i$ $\in$ I$_{wa}$: block size=watermark size **do**
  w$_a$ = (1/t)w-((1-t)/t)$\times$block$_i$: t=0.98 (2)
end loop
Output: attacked watermarks(w$_a$)

ducted on four 8-bits grayscale images (Lena, Barbara, Boat, and Cameraman) sized 512×512 pixels, and using watermark image sized 64×64. Initially, the targeted image is partitioned into 64×64 non-overlapping blocks, and the texture features for each block are analyzed to build the Boolean matrix. The Boolean matrix is used as input of Concept Explorer (ConExp) tool v1.3 (Yevtushenko et al. 2000), which provides basic functionality needed to extract the set of formal concepts. The high frequency objects (blocks), which frequently appear with most formal concepts, express the most textured blocks within the targeted image. These textured blocks are used as input for the watermark embedding process. The performance of the proposed watermarking system is evaluated in terms of perceptual quality and the robustness against different attacks. The PSNR, SSIM metrics are used to measure the perceptual quality of watermarked image with respect to the original image, while CC and BER metrics are used to measure the similarity and stabilization between the original watermark and the extracted watermark.

Perceptual quality analysis

The perceptual quality of embedded watermark in the host images is illustrated in figure 3. The mentioned PSNR and SSIM are computed without any image processing attack on watermarked images.Figure 3 shows that the PSNR for Lena, Barbara and Boat images reaches 44 dB, while for Cameraman image it is equal to 45.5 dB. On the other hand, the SSIM for Cameramen and Barbara image reaches 0.97 comparing with 0.93 for Lena and Boat images. The highly quality of embedded Cameraman image over other processed images comes from the lower number of blocks considered in watermark embedding, where the number of concerned blocks in Cameraman image is 18 blocks comparing with 30, 28, and 32 for Lena, Barbara, and Boat images respectively.

To ensure the efficiency of the proposed model in terms of perceptual quality, figure 4 presents PSNR comparison for Lena image between the proposed model and other models in (Lang et al. 2014, Han et al. 2016, kumar et al. 2016, Parah et al. 2016 and Ghadi et al. 2016). Some of these models such in (Lang et al. 2014, Han et al. 2016, kumar et al. 2016 and Parah et al. 2016) use the transformed coefficients to embed watermark, while other model such in (Ghadi et al. 2016) use the spatial domain.

Figure 4 shows that the achieved PSNR of the proposed model reaches 44 dB and outperforms the PSNR of other

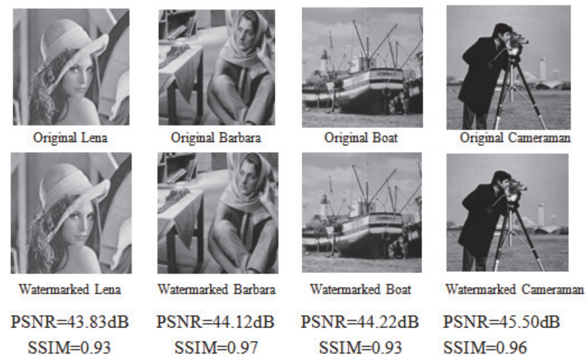

Figure 3: PSNR and SSIM of embedded images with respect to the original images.

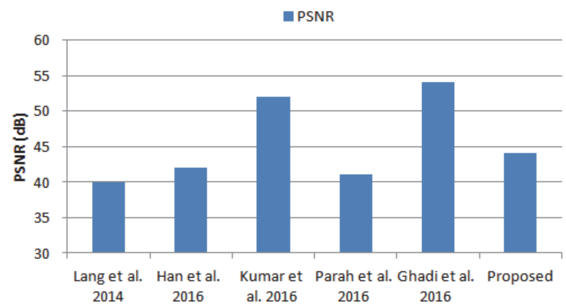

Figure 4: PSNR comparison for Lena image.

models in (Lang et al. 2014, Han et al. 2016 and Parah et al. 2016), where the PSNR ranged 40-41 dB. This result confirms the significance of texture features in defining more interesting blocks for watermark embedding with least degradation in perceptual quality of host image. On the other hand, the PSNR of the proposed model is lower than the PSNR of other models in (kumar et al. 2016 and Ghadi et al. 2016), where the PSNR is ranged 52-54 dB. This shortfall in the proposed model can be explained by the vital role of predefined metric (minimum support) in defining more appropriate blocks for watermark embedding in (Ghadi et al. 2016) model, as well by maintaining the indiscernible properties on the perceptual quality of host image in (kumar et al. 2016).

Robustness of watermarking system

To evaluate the robustness of the proposed system against different attacks. The embedded images are tested against common image-processing attacks including rotation, cropping, translation, JPEG compression, Gaussian noise, and median filtering. The presentation of these attacks is available in (Ghadi et al. 2017). These attacks are applied on embedded images using the StirMark tool v.4 (Petitcolas et al. 1998) and the Matlab R2016a. The experiments result for the four embedded images are convergent, therefore we only present the experiments result of Lena image.
Figure 5 presents the robustness of Lena image against different attacks in terms of CC and BER, which are computed

between the extracted watermark w_a and the original watermark w.

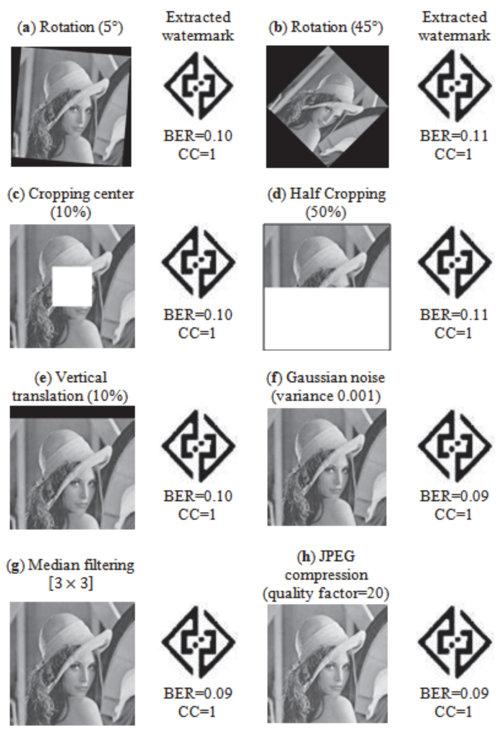

Figure 5: The robustness results for Lena image against geometric and non-geometric attacks.

From the mentioned results in figure 5, we can show that the CC in all cases equals 1, while the BER against non-geometric attacks such as (Gaussian noise, median filtering, and JPEG compression) is equal 0.09 and it is ranged 0.10-0.11 against geometric attacks such as (rotation, cropping, and translation). These results are expected due to the negative effect of geometric attacks on the watermarked image comparing with slight effect of non-geometric attacks. Moreover, the robustness of the proposed system against hybrid attacks is also tested. Figure 6 presents the robustness results for Lena image against two hybrid attacks.
The results in figure 6 ensure the robustness of the proposed

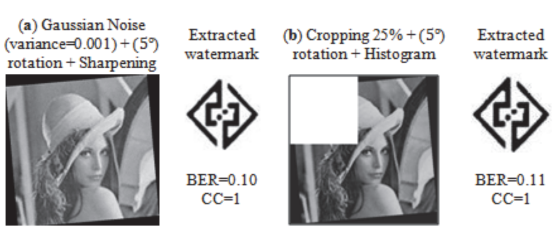

Figure 6: The robustness results for Lena image against hybrid attacks.

model against hybrid attacks, where the BER ranges 0.10-0.11 and the CC equals 1. Additionally, the results show that the robustness against hybrid geometric attacks is less than the robustness against non-geometric hybrid attacks.
To evaluate the performance of the proposed model, it has been compared with other watermarking models in terms of robustness. Figure 7 and figure 8 present BER comparison for Lena image after 5 degree rotation, 10% cropping from the center, Gaussian noise and median filtering between the proposed model and other models presented in (kumar et al. 2016, Parah et al. 2016 and Ghadi et al. 2016).

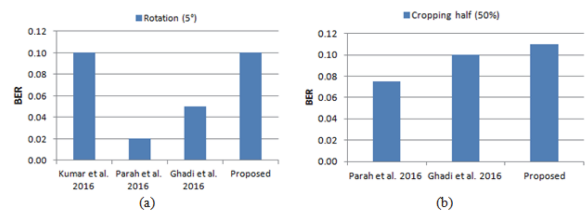

Figure 7: BER comparison for Lena image after 5 degree rotation and after 10% cropping from the center.

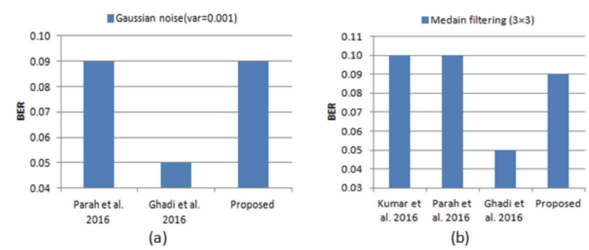

Figure 8: BER comparison for Lena image after Gaussian noise and median filtering.

From figure 7, we can show that the proposed model present an interesting BER against cropping attack comparing with (Parah et al. 2016), but it has a higher BER against rotation attack comparing with (Parah et al. 2016 and Ghadi et al. 2016). Subsequently, figure 8 shows that the achieved BER in this model is similar to the achieved BER in (kumar et al. 2016 and Parah et al. 2016) against Gaussian noise and median attacks, but it has a higher BER comparing with (Ghadi et al. 2016).

Capacity of watermarking system

To assess the capability of proposed watermarking system to accept large watermarks, we conduct a comparison between the performance of embedding 64×64 watermark and embedding 128×128 watermark under different attacks. This comparison is conducted on 512×512 (Barbara) image. The experiments result under different attacks is presented in figure 9.
The mentioned result in figure 9 shows that the CC ratio in all cases equals 1, while the BER with embedding 64×64 watermark is better than embedding 128×128 watermark by 2-3%. This can be explained by the amount of pixels that could be affected due to the attack on large size image.

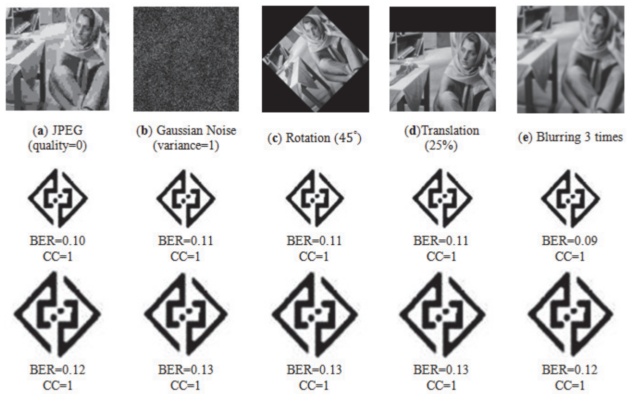

Figure 9: The robustness results of Barbara image against different attacks after embedding with 64×64 and 128×128 watermarks.

Time complexity

The execution time of the proposed model on processed images ranges 27-30 seconds. With note that all executions were achieved on Toshiba Satellite E45t-A Laptop/Intel core i5 2.3 GHz, 6.0 GB RAM. This time is reasonable and proves the efficiency of watermarking model to achieve image authentication across wireless networks.

Conclusion

This paper suggested a robust watermarking scheme to authenticate images across wireless networks based on FCA and texture analysis. FCA is used to find a meaningful knowledge that helps to embed the watermark efficiently, such as to obtain a robust watermarking. The formal concepts resulting from the application of the FCA method are exploited to extract highly textured blocks in the targeted image that are convenient with HVS and more preferable to embed the watermark robustly. The experiments result presented interesting ratios of imperceptibility and robustness against different attacks. The PSNR reached 45.5 dB, the BER ranged 0.9-0.11%, and the CC reached 1.
Future work include performance evaluation of the proposed model through real experiments on e-healthcare networks.

References

Sawant, R. 2010a. Image mining using texture and shape feature. *International Journal on Computer Science and Engineering*. 2(9): 2808-2815.

Priss, U. 2006. Formal concept analysis in information science. *Annual review of information science and technology*. 40(1): 521-543. ACM.

AlShaikh, M., et al. 2016a. Formal concept analysis to improve robustness on medical image watermarking schemes in the spatial domain. *International Journal of Computer Science and Information Security*. 14(2): 261-269.

Tilley, T., et al. eds. 2005. A survey of formal concept analysis support for software engineering activities. *Formal Concept Analysis*. 3626: 250-271. Springer.

Priss, U. eds. 2005. Linguistic applications of formal concept analysis. *Formal Concept Analysis* . 3626: 149-160. Springer.

Quist-Aphetsi, K. et al. 2015a. A spatial domain watermarking approach for digital images based on image features built on formal concepts analysis. *International Journal of Computer Applications*. 119(7): 5-10.

Ghadi, M., et al. 2016. A robust associative watermarking technique based on frequent pattern mining and texture analysis. *In Proceedings of the 8^{th} International Conference on Management of computational and collective IntElligence in Digital EcoSystems*. 73-81. ACM.

Poelmans J., et al. 2010. Formal concept analysis in knowledge discovery: a survey. *In the proceeding of the 18^{th} international conference on Conceptual structures: from information to intelligence*. 6208: 139-153. LNCS,: Springer.

Alqadah, F.; Bhatnagar, R. 2011a. Similarity measures in formal concept analysis. *Annals of Mathematics and Artificial Intelligence*. 61(3): 245-256. Springer.

Das, A. eds. 2015. Guide to signals and patterns in image processing: foundations, methods, and applications. *Springer International Publishing*. Switzerland.

Umbaugh, E-S. 2011a. Digital image processing and analysis: human and computer vision applications with CIVPtools. *Journal of Electronic Imaging*. 20(3): 335-399. CRC Press.

Yang, H.; Yin, J. 2013a. A secure removable visible watermarking for BTC compressed images. *Multimedia Tools and Application*. 74(6): 1725-1739. Springer.

Yevtushenko, S. 2000. System of data analysis ”Concept Explorer”. *In Proceedings of the 7^{th} national conference on Artificial Intelligence*. 127-134.

Lang, J., et al. 2014a. Blind digital watermarking method in the fractional Fourier transform domain. *Opt. Lasers Eng.*. 53: 112-121.

Han, J., et al. 2016a. A digital image watermarking method based on host image analysis and genetic algorithm. *J Ambient Intell Human Comput*. 7: 37-45.

Parah, S., et al. 2016a. Robust and blind watermarking technique in DCT domain using inter-block coefficient differencing. *Digital Signal Processing*. 53: 11-24.

Kumar, S., et al. 2016a. Rough set based effective technique of image watermarking. Journal of Computational Science. *Journal of Computational Science*. 1-17.

Ghadi, M., et al. 2017a. A novel zero-watermarking approach of medical images based on Jacobian matrix model. *Security Comm. Networks*. 1-16.

Petitcolas, F., et al. 1998. Attacks on copyright marking systems. *In Proceedings of 2^{nd} Workshop on Information Hiding*. 1525: 218-238. Springer.