THE JOURNAL OF DIGITAL FORENSICS, SECURITY AND LAW

Journal of Digital Forensics, Security and Law

Volume 12 | Number 4 Article 3

12-2017

Multiple Content Adaptive Intelligent Watermarking Schemes for the Protection of Blocks of a Document Image

Chetan KR Mr.
Jawaharlal Nehru National College of Engineering, chetankr@jnnce.ac.in

S Nirmala Dr.
Jawaharlal Nehru National College of Engineering, nir_shiv_2002@yahoo.co.in

MULTIPLE CONTENT ADAPTIVE INTELLIGENT WATERMARKING SCHEMES FOR THE PROTECTION OF BLOCKS OF A DOCUMENT

Chetan KR
Jawaharlal Nehru National College of Engineering

S Nirmala
Jawaharlal Nehru National College of Engineering

ABSTRACT

Most documents contain different types of information, such as white space, static information, and dynamic information, or mix of static and dynamic information. In this paper, multiple watermarking schemes are proposed for protection of the information content. The proposed approach comprises of three phases. In Phase-1, the edges of the source document image are extracted, and the edge image is decomposed into blocks of uniform size. In Phase-2, GLCM features like energy, homogeneity, contrast and correlation are extracted from each block and the blocks are classified as no-information, static, dynamic and mix of static and dynamic information content blocks. The adjacent blocks of same type are merged together into a single block. Each block is watermarked in Phase-3. The type and amount of watermarking applied is decided intelligently and adaptively based on the classification of the blocks which results in improving embedding capacity and reducing time complexity incurred during watermarking. Experiments are conducted exhaustively on all the images in the corpus. The experimental evaluations exhibit better classification of segments based on information content in the block. The proposed technique also outperforms the existing watermarking schemes on document images in terms of robustness, accuracy of tamper detection, and recovery.

1. INTRODUCTION

Document images are used as means of identification and ownership in various transactions performed online [1-3]. For instance, the cheque image indicate the ownership of the account holder in business transactions, marks cards identify the result of a particular student. Since most of transactions use digital data, there is a growing need of security in these transactions. From past two decades, watermarking is used as a primary means of authentication and protecting intellectual properties. Further, the document image watermarking is an active research area as it offers less redundancy for watermarking and semantics of information content should not change after watermarking. The nature and structure of the document image also brings new challenges. The document image

comprises of blocks with different levels of sensitivity. For instance, in a cheque image, there is a lot of non-information blocks, blocks containing static information like bank name, preprinted information, segments containing information like date, which is a mix of dynamic information and some static information like "/" symbol and some blocks contain dynamic information such as signature of the account holder. Each of the blocks possess different levels of sensitivity and requires different levels of protection. This could be achieved using multiple watermarking schemes for blocks of document image. For blocks with no information, watermarking is not required. Robust watermarking can be applied on the blocks of a document image that only contain static information. For blocks with only dynamic information, fragile watermarking technique can be used. For a block of a document image that contains mix of static and dynamic information content, semi-fragile watermarking mechanism can be applied. In this paper, appropriate watermarking scheme (robust/fragile/semi-fragile) is applied based on the type of block classified is applied. Thus, multiple watermarking is needed for effective authentication of the document images. Multiple watermarking schemes fulfil many objectives: improve the perceptual quality of the watermarked image by reducing embedding capacity; perform tamper detection and recovery with better accuracy; incur lesser time as only few blocks need to be watermarked and less embedding capacity is sufficient.

This paper is organized as follows: Section 2 provides a literature review of the multiple watermarking schemes. The proposed model is explored in Section 3. Section 4 presents experimental results of the proposed multiple watermarking scheme. Performance analysis of the results are detailed in Section 5. Conclusions of the proposed work are summarized in the last section.

2. LITERATURE REVIEW

The watermarking algorithms can be classified as robust, fragile and semi-fragile based on the robustness to incidental and intentional attacks [4]. The prominent efforts in area of robust, fragile and semi-fragile watermarking techniques on document images can be found in references 5-10. Amir et al., [11] proposed a watermarking technique based on the entropy masking feature of the Human Visual System (HVS). Kankanhalli et al. [12] developed a watermarking technique by embedding just noticeable watermarks. Radharani et al. [13] designed a content based watermarking scheme in which watermark is generated using Independent Component Analysis (ICA) for each block of the input image. In [14-16], the segmentation of the image into objects using image statistics and subsequently applying the robust watermarking schemes for each object is discussed. Chin-Shiuh Shieh et al. [17] proposed the use of genetic algorithm (GA) [18] to compute the optimal frequency bands for watermark embedding into a Discrete Cosine Transform (DCT) based watermarking system. This system can simultaneously improve security, robustness, and image quality of the watermarked image.

All these existing schemes apply a single watermarking technique on the entire image. In the recent past, many works on multiple watermarking has been reported. Sameh et al. design a watermarking scheme that embeds multiple watermarks with different compression domains into the same source [19]. The watermark was compressed in Discrete Cosine Transform, Discrete Wavelet Transform and Singular Value Decomposition domains and multiple copies of the watermark is embedded into the source image using Least

Significant Bit substitution technique. This multiple watermarking scheme was successful in providing robustness to many image processing attacks. Zhe-Ming et al., [20] developed an algorithm for embedding multiple watermarks into the Vector Quantization (VQ) domain, as well as for hiding the secret keys associated with the watermarks in the transform domain to enhance the robustness of the watermarked image. The watermarks were embedded to satisfy goals of both content authentication and copyright protection. Semi-fragility and robustness were exhibited by the multi-purpose watermarking scheme [20]. Nicholas Paul et al., [21] explored the different ways of multiple watermarking like rewatermarking, segmented watermarking and composite watermarking. They described different attack scenarios [22] and level of robustness that could be provided by each category of multiple watermarking. A novel method that involves watermarking multiple copies of biometric fingerprints using texture properties of wavelet transforms was put forth in reference 23. Two level Haar Wavelet was applied on the biometric fingerprints and texture analysis was performed [23]. The texture properties like entropy and homogeneity were used to decide the number of watermarks. The watermarks were further secured using the concept of Visual Cryptography [24].

The literature reviews on the content based multiple watermarking techniques reveals that most of the existing works does not perform intelligent classification of blocks of a document image based on the nature of information in the document. In the existing techniques, multiple watermarks of the same type are embedded to each block of the document image. In addition, the existing schemes also incur tradeoff between robustness and fragility of the watermarking multiple times. A novel multiple watermarking scheme to address these issues was discussed in [25]. In this work, the document image is decomposed into blocks of size 128 X 128. The blocks are classified based on the sensitivity level of the information present in the block. Next, fragile or robust watermarking technique is applied based on the protection level required by the information content in the block. Experimental results [25] reveal that classification of the block based on information content in the block is achieved with good accuracy (93%). However, there are some regions in which blocks are incorrectly classified due to fixed size of the block.

In this work, the division of the image into blocks such that adjacent regions with same type of information content are grouped together is proposed. The type and amount of watermarking is decided intelligently and adaptively based on the information content in the block. This results in less embedding capacity for each block and subsequently improves fidelity of the watermarked image. The time required for embedding and extraction of watermarks is reduced due to the application of required level and type of watermarking for each block. This makes proposed scheme perform better than the existing multiple watermarking scheme [25].

3. PROPOSED MULTIPLE WATERMARKNG MODEL

The proposed multiple watermarking model consists of two processes namely Embedding and Extraction. The details of embedding and extraction process are discussed in the following subsections.

3.1 Embedding Process of multiple watermarks

The embedding process performs multiple intelligent watermarking based on the information content in the image. There are three stages involved during the embedding of watermarks in the proposed scheme, namely, Formation of Blocks, Fusion of Blocks and Multiple Watermarking. These stages are depicted in Figure 1.

3.1.1 Formation of Blocks

In this stage, the input color document image is converted to grayscale using color conversion from RGB to YC_bC_r [29]. Luminance channel (Y-component) consists of the energy of an image and is thus a primary carrier of information content of an image. The decision on the type of watermarking required depends on the information content in an image and thus only luminance channel is used for watermarking process. The edges of the luminance component of the document image is detected using Canny edge detector. The reason for selecting Canny edge operator is that it exhibits better edge detection capabilities in presence of noise and is also able to detect weak edges [30]. The edge image is divided into blocks of size 32 X 32. The blocks are set to the size 32 X 32. The rationale behind setting of this block size is a balance accuracy of tamper classification and processing time as in references 26-28. The pixels in each block contain gray levels with values in the range 0-255.

3.1.2 Fusion of Blocks

The texture of each block in the edge image is analyzed using statistical features of Gray Level Co-occurrence Matrix (GLCM). It defines both the structural as well as spatial properties of an image texture [32].

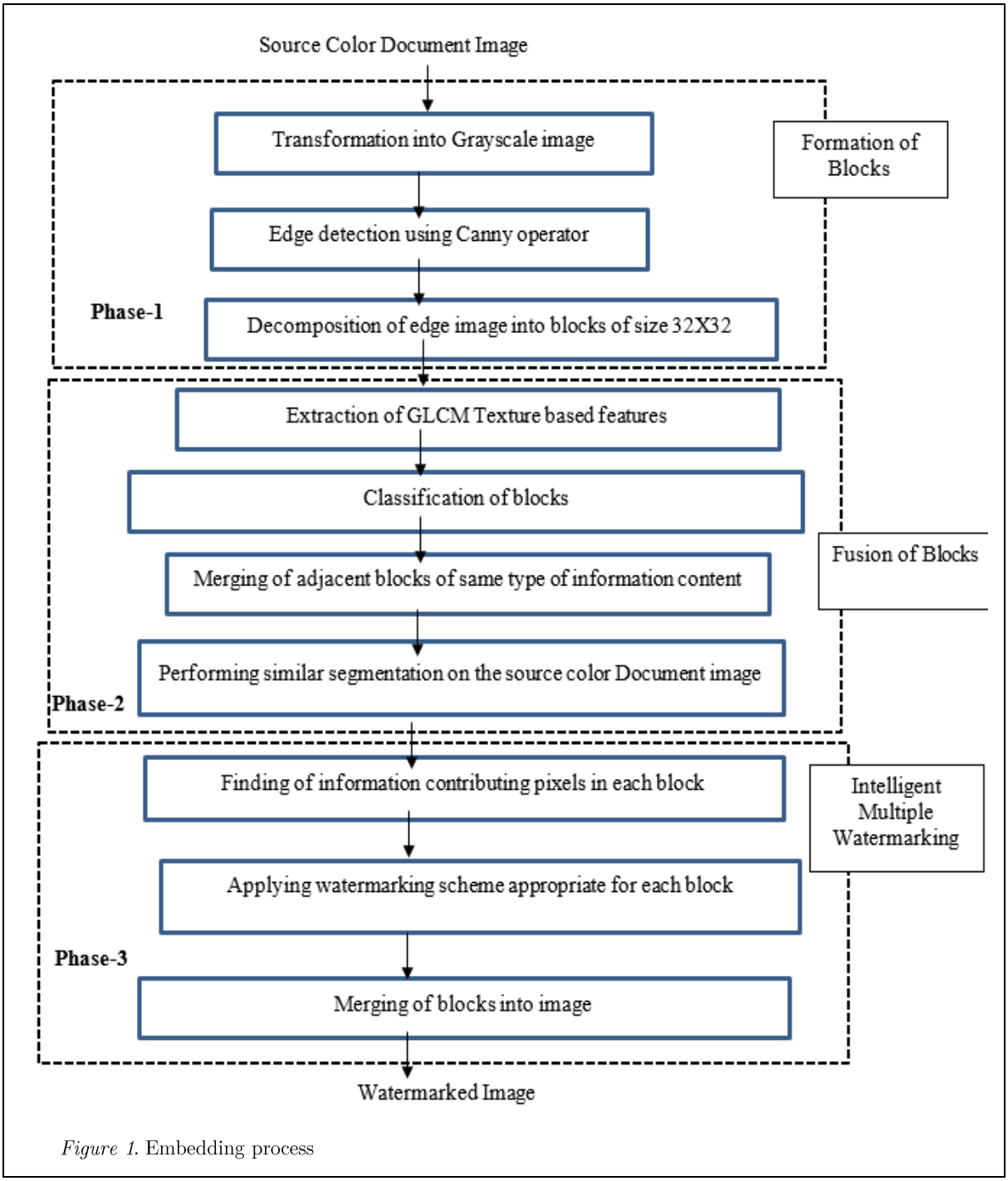

Figure 1. Embedding process

In this work, GLCM features like Energy, Homogeneity, Contrast, and Information Measure of Correlation are extracted for each image block using unit distance and in four directions i.e. {0°, 90°, 180°, 270°}. The features are computed as follows:

(i) Energy –Energy of the block of a document image can be computed using equation below:

$$\boldsymbol{Energy(b) = \sum_{i=1}^{32}\sum_{j=1}^{32} G_{ij}^{2}} \quad (1)$$

where b is the block of size 32 X 32, G_{ij} represents the GLCM value of the pixel at the location *(i,j)*. All the gray level pairs i and j in the image block are used to compute the energy of that block.

(ii) Homogeneity –Homogeneity of a block of document image is measured using the following equation:

$$\boldsymbol{Homogeneity(b) = \sum_{i=1}^{32}\sum_{j=1}^{32} \frac{1}{1+(i-j)^2} G_{ij}} \quad (2)$$

(iii) Contrast –The contrast present in the block of a document image is computed using equation as follows:

$$\boldsymbol{Contrast(b) = \sum_{i=1}^{32}\sum_{j=1}^{32} (i-j)^2 G_{ij}} \quad (3)$$

(iv) Information Measure of Correlation –It is computed as an average of two information measures of correlation (IC1 and IC2), which are given by the equations below:

$$\boldsymbol{IC1(b) = \frac{H-H1}{\max(H,H1)}} \quad (4)$$

$$\boldsymbol{IC2(b) = \sqrt{1-\exp(-2(H2-H))}} \quad (5)$$

where H, H1 and H2 are given by:

$$\mathbf{H} = -\sum_{i=1}^{32}\sum_{j=1}^{32} G_{ij} \log_2 G_{ij} \quad (6)$$

$$\mathbf{H1} = -\sum_{i=1}^{32}\sum_{j=1}^{32} G_{ij} \log_2 (G_x(i) G_y(j)) \quad (7)$$

$$\mathbf{H2} = -\sum_{i=1}^{32}\sum_{j=1}^{32} G_x(i) G_y(j) \log_2 (G_x(i) G_y(j)) \quad (8)$$

where, $G_x(i)$ is a marginal probability distribution obtained by row-wise sum of GLCM values for the block of the document image.

These four GLCM features can be used to intelligently classify the type of the block based on the information content in the block. The blocks can be categorized as: no information, static, dynamic, or mixture of static and dynamic. No information blocks have either 'no information content' or 'no variation in the information content.' Static blocks have minimal distribution in their information content and have more correlation in adjacent pixels. Dynamic blocks contain more disorders in their information content and are less correlated with moderate energy. Thus, all four parameters are required for better classification of the blocks in the document image. Energy parameter can measure amount of information content, homogeneity measures global variations and contrast measures the local variations in the information content and correlation measures the linear dependencies in the information content in each block of the document image.

From experimental evaluations, a range of values is set as threshold for all the parameters as shown in Table 1 for identification of type of the block intelligently. It is observed from the threshold values in Table 1 that the parameters exhibit overlapping values and are all required to make a final decision on the classification of the type of the block.

Table 1.
Classification of the types of the block

Energy	**Homogeneity**	**Contrast**	**Information measure of Correlation**	**Type of the block**	**Type of the watermarking scheme**
0-0.4	0.7-1.0	0-0.19	0.9-1.0	No information	No
0.11-0.5	0.6-0.9	0.1-0.4	0.8-1.0	Static	Robust
>0.5	0.41-.79	>0.3	0.5-.79	Mixture of static and dynamic	Semi-fragile
>0.5	0.0-0.5	>0.35	0.1-0.6	Dynamic	Fragile

As discussed in the Introduction section, each type of the block of document image needs different levels of security and subsequently a different type of watermarking. The type of watermarking required for each category of the block is shown in Table 1.

The accuracy in the identification of the type of the block is measured by percentage of match between expected and identified type of the block. The expected type is labeled by a human expert. The identified type is computed automatically using classification criterion in Table 1. Table 2 shows the accuracy of identification of the type of the block with GLCM parameters. The accuracy values in Table 2 reveals that accuracy of identification is high when all the four parameters are used.

A sample cheque document image (input image) is shown in Figure 2. The classification of the type of the blocks for this sample Cheque document image is shown in Figure 3. It can be observed from Figure 3 that blocks containing uniform texture of less information content (like background patterns in the cheque) has been marked as static blocks. The blocks having lot of variations in the information content like Signature; Cheque numbers are identified as dynamic blocks. Some blocks containing preprinted textual information are labeled as mixture of static and dynamic blocks. Unmarked regions in the document image are no-information blocks where watermarking is not required.

Table 2.
Accuracy in the identification of the types of the block

Parameters used	**Accuracy of identification (in %)**
Energy, Homogeneity, Contrast, Information measure of Correlation	98.2
Energy, Homogeneity, Contrast	93.7
Energy, Homogeneity, Information measure of Correlation	88.7
Energy, Contrast, Information measure of Correlation	73.3
Homogeneity, Contrast, Information measure of Correlation	59.1

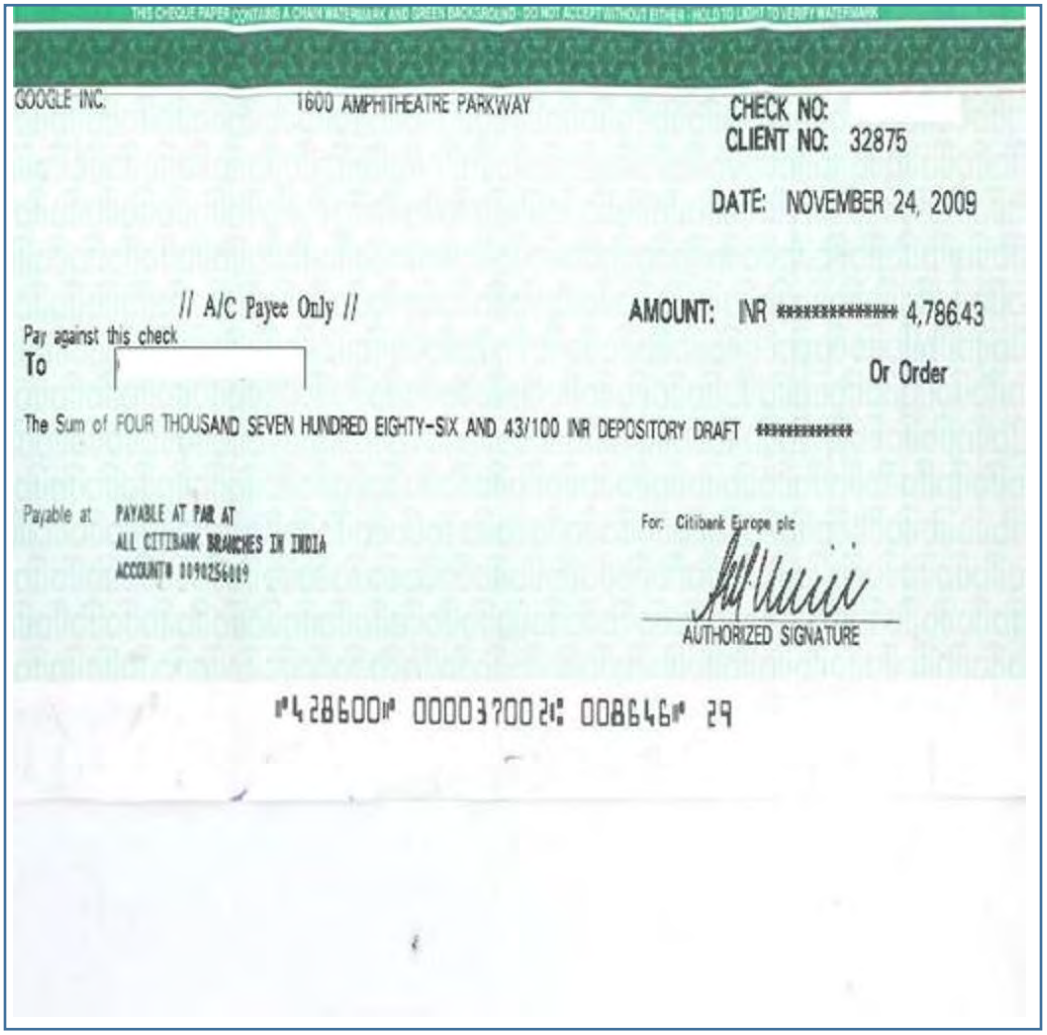

Figure 2. Classification of the blocks of sample Cheque image

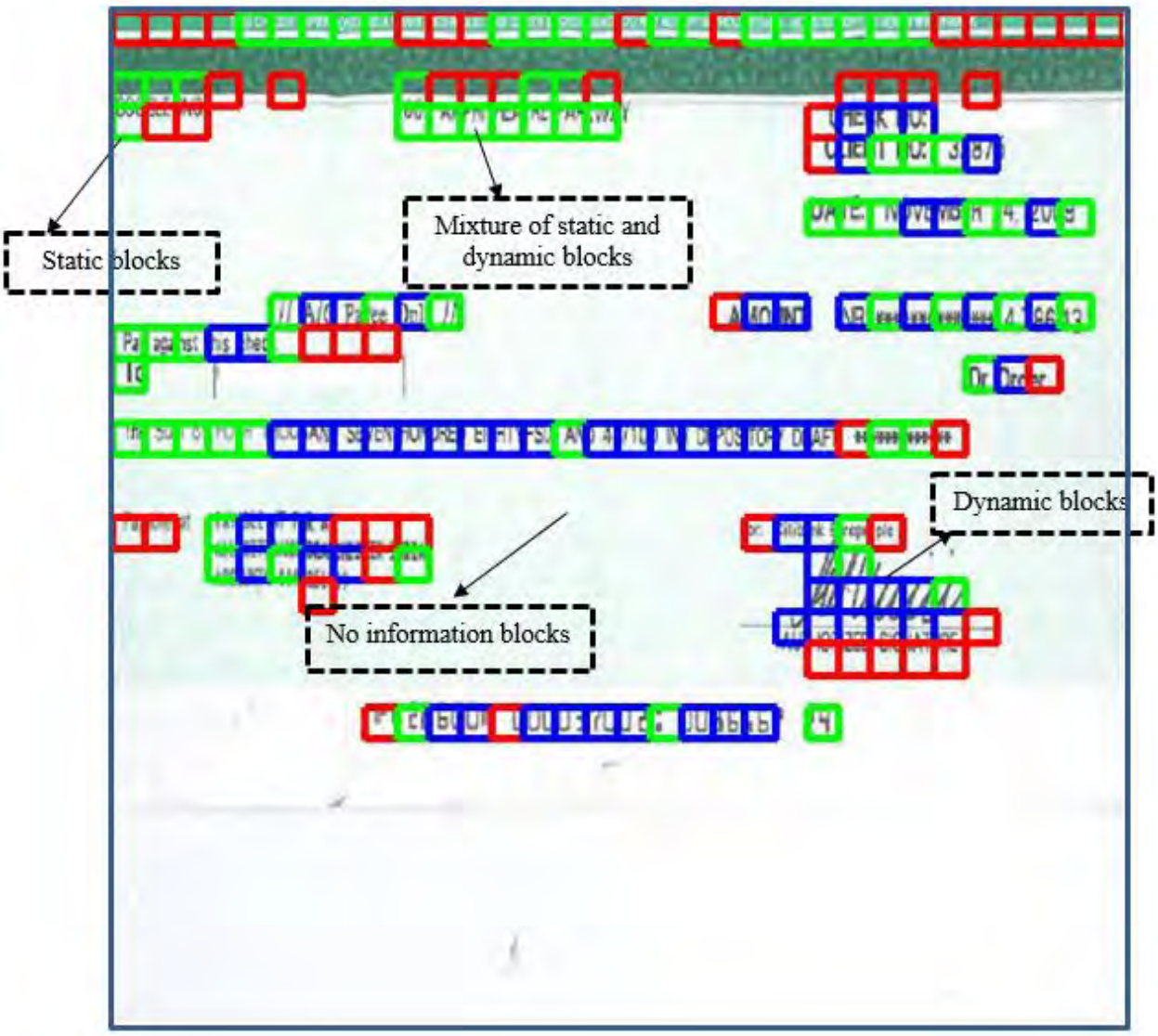

Figure 3. Classification of the blocks of sample Cheque image

Once all the blocks of a document image are classified into appropriate types, adjacent blocks which are of the same type are merged together into a new block. This reduces the amount of time required for watermarking. The blocks formed after merging adjacent blocks of same type are shown in Figure 4. It can be revealed from Figure 4 that most of the blocks of same type are adjacent to each other, and segments formed represents the objects of the document image. Subsequently for each block, only the pixels contributing information is found by using only those pixels of the corresponding edge image with pixel values 1. This further improves perceptual quality of the watermarked image and helps in localization of tampering. The pixels used for watermarking the blocks in Figure 4 is depicted in Figure 5. It can be observed from Figure 5 that there exists lot of non-information content regions within a block.

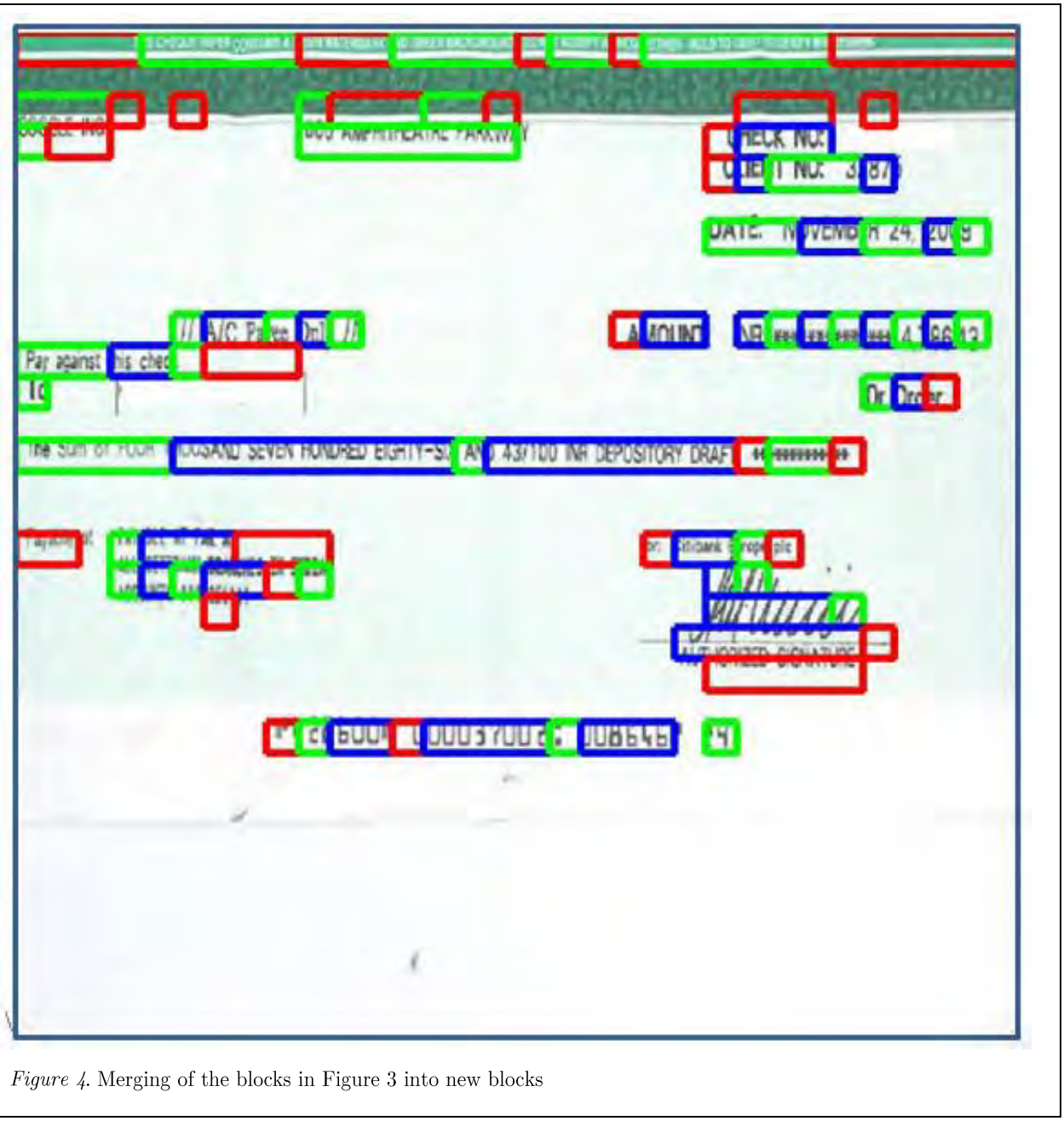

Figure 4. Merging of the blocks in Figure 3 into new blocks

3.1.3 Intelligent Multiple Watermarking

The type of watermarking is decided by the type of the block as given in Table 1.

3.1.3.1 Robust Watermarking

The watermarkable pixels in the blocks formed by static blocks are stored in a one-dimensional vector. The vector is protected by one-dimensional variant of the existing Integer Wavelet based Robust watermarking [26]. The algorithm for robust watermark embedding is given below:

Algorithm: One-dimensional Integer Wavelet Robust Watermarking

Step 1: A binary image is selected as the watermark.

Step 2: The watermark is encoded using binary coding technique in reference 26.

Step 3: The one-dimensional vector of watermarkable pixels are transformed using one-dimensional Integer Haar Wavelets for level-2.

Step 4: Watermark bits are redundantly embedded into second level approximate coefficients using quantization-based embedding [26].

Step 5: The wavelet coefficients are inverse transformed to obtain Watermarked block.

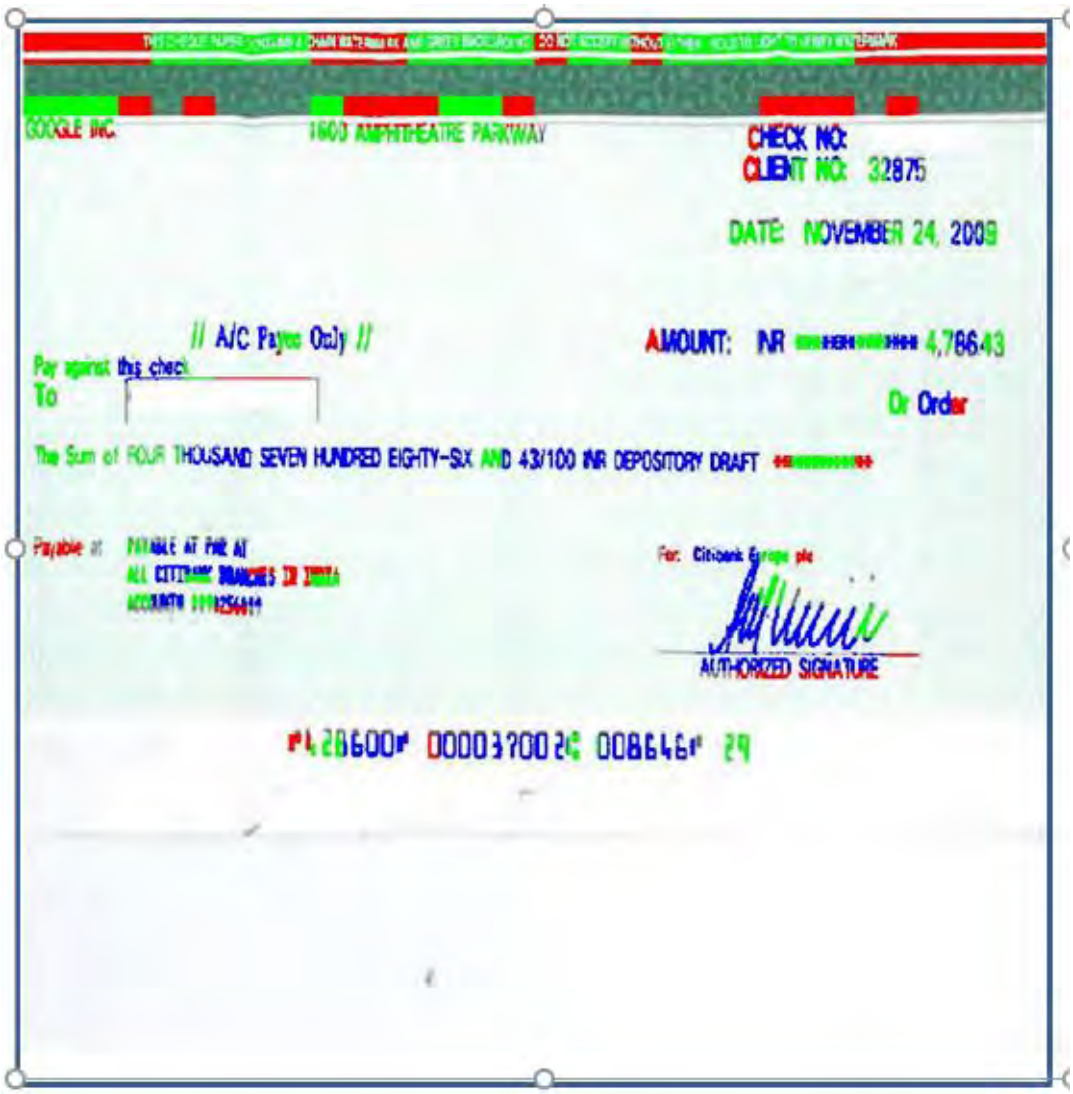

Figure 5. Watermarkable pixels of each block

3.1.3.2 Semi-Fragile Watermarking

The watermarkable pixels in the blocks formed by blocks whose type is mixture of static and dynamic information content is vectorized into one-dimension and the vector is protected using semi-fragile watermarking. In the proposed work, the semi-fragile watermarking is implemented using an existing technique based on curvelets modified for one-dimension [27]. The algorithm for curvelet-based embedding is as follows:

Algorithm: One-dimensional Curvelet based semi-fragile watermark embedding

Step 1: The one-dimensional vector of watermarkable pixels is transformed using Discrete Curvelet Transformation [27].

Step 2: The first level coarse curvelet coefficients are quantized into 4 bits.

Step 3: Locations for embedding is determined intelligently using content adaptive technique [27].

Step 4: Embedding is done using replacement of four Least Significant Bits of first level coarse curvelet coefficients.

Step 5: Watermarked block is obtained after applying one-dimensional inverse Curvelet Transform.

3.1.3.3 Fragile Watermarking

The watermarkable pixels in the blocks formed by blocks containing dynamic information content are secured using a fragile watermarking technique. The pixels in the blocks are reshaped into one-dimensional vector and are protected by modifying an existing fragile watermarking technique based on contourlets for one-dimension [28]. The procedure for fragile watermark embedding is given by the algorithm below:

Algorithm: One-dimensional Contourlet based fragile watermark embedding

Step 1: The one-dimensional vector of watermarkable pixels are transformed using Discrete Contourlet Transformation [28].

Step 2: The first and second level contourlet coefficients are used to generate watermark.

Step 3: Locations for embedding is determined intelligently using content adaptive technique [27].

Step 4: Embedding is done using quantization of contourlet coefficients [28].

Step 5: The contourlet coefficients are inversely transformed to obtain watermarked block.

All the watermarked blocks are merged together to obtain a watermarked document image.

3.2 Extraction Process of multiple watermarks

The extraction process is applied on the watermarked image. The extraction process is similar to the embedding process, until the finding of contributing pixels in the blocks of the watermarked image. The watermarks are then extracted based on the type of the blocks forming the blocks of the watermarked image.

3.2.1 Robust Watermark Extraction

For the blocks containing static information content, the following robust watermark extraction algorithm is applied:

Algorithm: One-dimensional Integer Wavelet based Robust Watermark Extraction

Step 1: The one-dimensional vector of watermarkable pixels in the blocks of the watermarked image is transformed using one-dimensional Integer Haar Wavelet of level-2.

Step 2: The encoded watermark bitstream is extracted from level-2 approximate coefficients [26].

Step 3: The encoded bitstream is decoded using binary decoding technique [26].

Step 4: All multiple redundant copies are extracted

Step 5: A majority voting scheme [26] based watermark is formed using all the extracted watermarks.

Step 6: The watermark is compared with the embedded binary image and authentication decision is made.

3.2.2 Semi-fragile Watermark Extraction

Semi-fragile watermark extraction process is applied for the blocks formed by the blocks containing mixture of static and dynamic information content. The different steps involved in semi-fragile watermark extraction based on One-dimensional curvelets [27] is given below:

Algorithm: One-dimensional Curvelet based semi-fragile watermark extraction

Step 1: The one-dimensional vector of watermarkable pixels in the blocks of the watermarked image is transformed using Discrete Curvelet Transformation [27].

Step 2: The first level coarse curvelet coefficients are quantized into 4 bits.

Step 3: The embedded coefficients are extracted intelligently using content adaptive technique [27].

Step 4: The generated and extracted curvelet coefficients are compared using Feature Similarity Index [27].

Step 5: A tamper matrix is formed which shows the locations containing mismatches in the generated and extracted curvelet coefficients.

3.2.3 Fragile Watermark Extraction

For the blocks formed by the blocks containing dynamic information content, fragile watermark extraction algorithm is applied. The procedure for fragile watermark extraction using contourlets [28] is outlined below:

Algorithm: One-dimensional Contourlet based Fragile Watermark Extraction

Step 1: The one-dimensional vector of watermarkable pixels is transformed using Discrete Contourlet Transformation [28].

Step 2: The first and second level contourlet coefficients are used to generate watermark.

Step 3: The contourlet coefficients embedded in the locations determined intelligently using content adaptive technique [27] are extracted.

Step 4: The generated and extracted contourlet coefficients are compared using Feature and Structure Similarity Index (FSSI) [28].

Step 5: If there is a mismatch, tamper recovery is performed by replacing the embedded contourlet coefficients into generated coefficients [28].

4. EXPERIMENTAL RESULTS

We have created a corpus of document images. All the images in the corpus are scanned document images. The classes of document image in the corpus are cheques, receipts, identity cards, marks cards and certificates. There are 300 document images in the corpus with 60 images for each class of document image. The watermark embedding and extraction has been tested for all the images in the corpus. The results of multiple watermark embedding and extraction for a sample document image is shown in Figure 6. It can be observed from the embedding result in Figure 6(c) that watermarked image is perceptually similar to the original document image. The watermarked image is subjected to an incidental attack namely salt and pepper noise, of 5%. From the robust watermark extraction result in Figure 6(d), it is evident that extracted logo is visually similar to logo used during embedding. The watermarked images are also subjected to intentional attacks like deletion, and modification of the information content in the watermarked image. It can be revealed from the visual inspection of the results in Figure 6(h), that the localization and detection of tampered information content exhibited by the proposed semi-fragile watermark extraction is accurate. Better recovery of tampered information contents

using the proposed fragile watermark extraction is also visible in the Figure 5(i).

5. PERFORMANCE ANALYSIS

The performance of the proposed multiple watermarking system is measured in terms of the following parameters: (i) accuracy in the intelligent classification of the type of the blocks of the document image, (ii) fidelity analysis using Peak Signal to Noise Ratio (PSNR), (iii) profiling of runtime, (iv) robustness analysis using Normalized Correlation Coefficient (NCC), (v) accuracy of tamper detection, and (vi) accuracy of tamper recovery.

Figure 6. Results of proposed multiple watermarking system (a) Source document image (b) Original watermark for robust watermarking (c) Watermarked image (d) Zoomed up segment containing static information content subjected to 5% Salt and Pepper Noise (e) Zoomed up segment containing mixture of static and dynamic information content subjected to deletion attack (f) Zoomed up segment containing dynamic information content being tampered with a forged signature (g) Extracted robust watermark (h) Tamper detection results (i) Tamper recovery result

5.1 Accuracy in the Intelligent Classification of the type of the blocks of a document image

The accuracy in intelligent classification of the type of the blocks of the proposed multiple watermarking system is compared with the existing system [25] is detailed in Section 2. The comparison has been made for different classes of the document image in the corpus and results of accuracy is tabulated in Table 3. It could be inferred from the accuracy values in Table 3 that proposed method exhibits better intelligence in classification of the types of the blocks of document image than the existing method [25]. The advantage of this more accurate classification in turn helps in appropriate type and amount of watermarking. This significantly improves the perceptual quality of the watermarked image and enhances the accuracy of tamper detection and recovery. The perceptual quality of the watermarked image is measured in terms of PSNR.

Table 3.
Accuracy in the Intelligent Classification of the blocks

Class of the Document Image	**Accuracy of identification (in %)**	
	Existing method [25]	**Proposed Method**
Cheques	93.42	97.92
Identification Cards	92.97	98.11
Marks Cards	93.11	97.87
Receipts	94.08	98.41
Certificates	91.96	98.31

5.2 Comparative Analysis

The performance of the proposed scheme is compared with existing methods [25-28]. The proposed scheme contains multiple watermarking techniques, such as robust, fragile and semi-fragile. The goals of each technique are different: robust watermarking aims at achieving robustness against all incidental attacks, fragile watermarking at tamper recovery and semi-fragile at tamper detection and localization. Hence metrics used for each technique is different. Normalized Correlation Coefficient (NCC) is used to measure robustness, Tamper Recovery Accuracy (TRA) is used to evaluate accuracy of tamper recovery and Tamper Detection Accuracy (TDA) is used to measure tamper detection and localization. These are elaborated in following subsections:

5.2.1 Fidelity Analysis

The fidelity of the proposed multiple watermarking scheme is evaluated in terms of PSNR. The perceptual quality of the watermarked image of size NXN is measured using Peak Signal to Noise Ratio (PSNR) [33]. A graph of PSNR values is depicted in Figure 7 for different classes of the document images.

The graph shown in Figure 7 reveals that PSNR values of the multiple watermarking schemes are better than robust [26] and fragile watermarking [28] schemes when applied separately. It could also be observed from PSNR values in Figure 7 that fidelity of the proposed multiple watermarking scheme is better than the existing multiple watermarking scheme [25]. This increase in PSNR and subsequently the perceptual quality of the watermarked image is due to two reasons: (i) More accuracy in the classification of the

blocks, and (ii) Grouping of similar adjacent blocks into new block. The quantity of the watermark to be embedded depends on the type of the block. Hence, the noise induced due to watermarking is reduced to a good extent and this results in the better fidelity of the watermarked image.

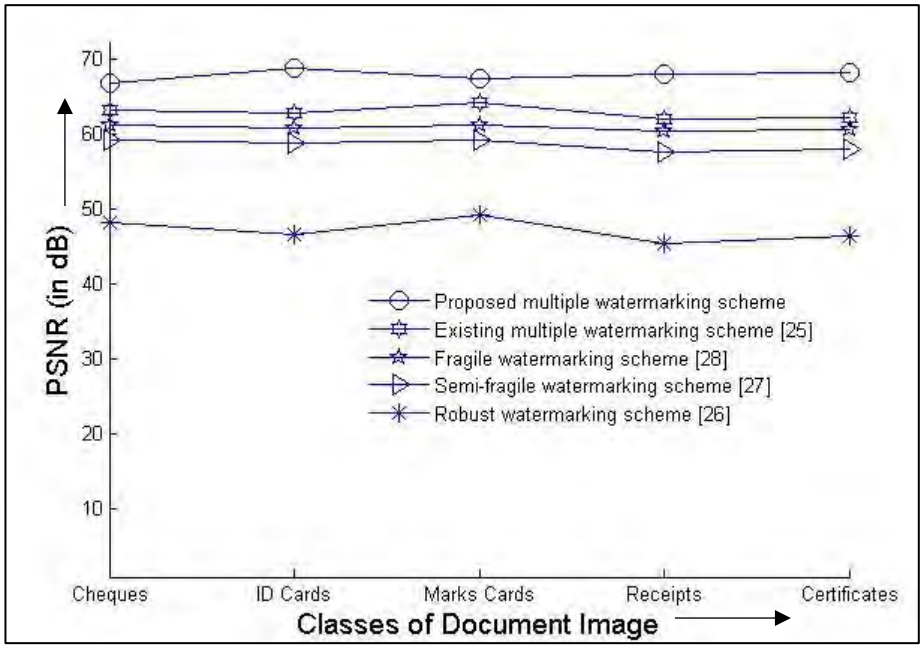

Figure 7. Effect of watermarking schemes on PSNR values of different classes of document images in the corpus

5.2.2 **Robustness Analysis**

The robustness capability of the proposed multiple watermarking scheme is evaluated by applying various attacks such as horizontal cropping, vertical cropping, resizing, noise and JPEG compression on all the document images in the corpus. The degree of robustness obtained is evaluated in terms of NCC: [34]. A comparative study of average values of NCC of the proposed and existing methods [25,26] is carried out and analysis is shown in Table 4. The NCC values in Table 4 show that proposed multiple watermarking scheme exhibits a slight improvement in the robustness of the watermarked image compared to the existing methods [25, 26]. The increase in robustness is due to the localization of robustness to the blocks that contain static information content.

Table 4
Average NCC values for different incidental attacks

Incidental Attack	**Existing Robust Watermarking scheme [26]**	**Existing Multiple Watermarking scheme [25]**	**Proposed Multiple Watermarking Scheme**
Salt and Pepper Noise	0.93	0.96	0.98
Cropping	0.97	0.97	0.98
Resizing	0.94	0.95	0.97
JPEG Compression	0.94	0.96	0.99

5.2.3 Accuracy of Tamper detection

Accuracy of tamper detection is evaluated as follows:

$$\boldsymbol{TDA = 1 - \frac{\sum_{i=1}^{n}(ta_i \oplus td_i)}{n}} \quad \textbf{(9)}$$

where, n – total number of bits in the semi-fragile watermarked blocks, ta – tampered bit, td – tamper detection bit. The TDA values are affected due to false positive and false negative alarms in the detection of tampered information content. Different intentional attacks like the insertion, deletion, modification and combination of them is applied on the watermarked image and the average values of TDA are recorded for proposed and existing methods [25, 27]. Figure 8 shows the graph that depicts the average values of TDA vs various intentional attacks. It can be observed from graph in Figure 8, that the proposed multiple watermarking schemes outperforms compared to existing methods [25,27] in the detection of tampered information content of document image.

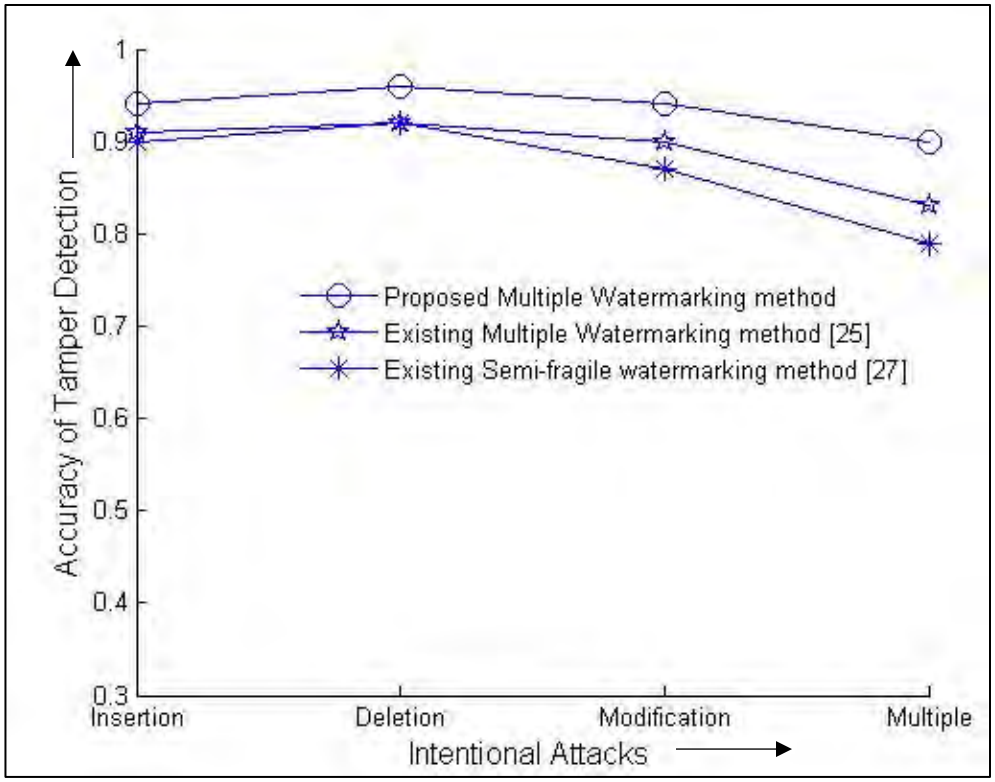

Figure 8. Average TDA values for different Intentional Attacks

5.2.4 Accuracy of Tamper recovery

Accuracy of tamper recovery is computed using the equation below:

$$\boldsymbol{TRA = 1 - \frac{\sum_{i=1}^{n}(ta_i \oplus tr_i)}{n}} \qquad \textbf{(10)}$$

where, n – total number of bits in the fragile watermarked blocks, ta – tampered bit, tr – tamper recovered bit. The TRA values are affected by false positive and false negative alarms in the recovery of tampered information content. The average values of TRA is computed for all document images in the corpus under different intentional attacks for proposed multiple watermarking scheme, existing multiple watermarking scheme [25] and an existing fragile watermarking method [28]. A graph is plotted that depicts the average values of TRA and is shown in Figure 9. It is evident from the graph in Figure 9, that the proposed multiple watermarking schemes performs better recovery of the tampered information content than the existing methods [25, 28].

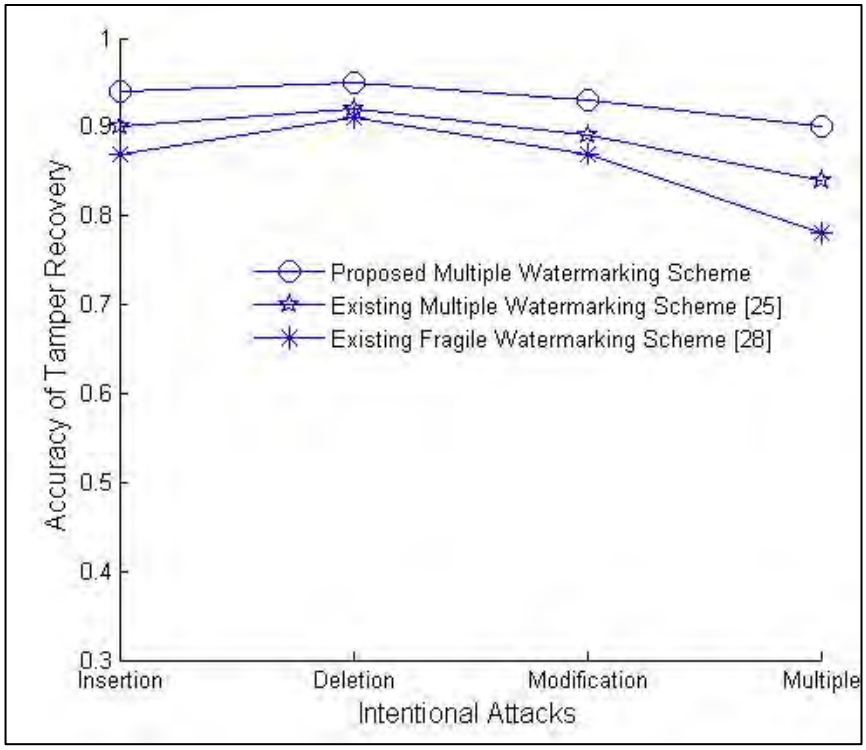

Figure 9. Average TRA values for different Intentional Attacks

5.3 Profiling of runtime

The watermarking system developed has been tested on a machine with Intel Core I3 processor and 2GB RAM. The average execution time (in seconds) for embedding and extraction for the proposed multiple watermarking system, existing multiple watermarking system [25], and existing methods [26-28]. The results of profiling of runtime is shown in Table 5. It can be observed that there is a considerable reduction in the execution time during embedding and extraction in the proposed work than existing methods [25-28]. As embedding and extraction is done only in the watermarkable pixels of the segments of the document image and amount of watermarking depends on the information content in the block of the document image, we achieved a substantial reduction in run time of proposed method.

Table 5
Average Execution Time

Method	**Embedding Time (in secs)**	**Extraction Time (in secs)**	**Total time (in secs)**
Existing robust watermarking method [26]	5.64	4.64	10.28
Existing semi-fragile watermarking method [26]	7.12	6.87	13.99
Existing fragile watermarking method [27]	7.32	7.04	14.36
Existing multiple watermarking method [25]	6.1	5.34	11.44
Proposed method	3.9	3.41	7.31

6. CONCLUSIONS

In this paper, an intelligent classification of the blocks of the document image according to the information content is performed. A novel framework for multiple watermarking schemes adaptive to the information content in the segment is proposed in this paper. The performance analysis of the proposed approach reveals improvement in the perceptual quality of the watermarked image. The proposed scheme is computationally less expensive than the existing method [25]. The proposed scheme also outperforms compared to the existing methods [25-28] in providing robustness, tamper detection and recovery capabilities. Adding intelligence in classification of intentional attacks is considered as future study of the current work.

REFERENCES

[1] M. Wu and B. Liu, (1998), Watermarking for Image Authentication, Proc. Of the IEEE Int. Conf. On Image Processing, pp. 437-441.

[2] Ingemar Cox , Matthew Miller, Jeffrey Bloom, Jessica Fridrich and Ton Kalker, (2007), Digital Watermarking And Steganography, Morgan Kaufmann Publishers Inc. San Francisco

[3] F Hartung and M Kutter, (2002), Multimedia Watermarking Techniques, Proceedings of IEEE, Vol. 87. No 7, pp. 1079-1107

[4] Potdar, V.M.; Song Han; Chang, E., (2005), A Survey of Digital Image Watermarking Techniques, 3rd IEEE International Conference on Industrial Informatics, pp.709, 716, doi: 10.1109/Indin.2005.1560462.

[5] H. Mirza, H. Thai, And Z. Nakao, (2008), Color Image Watermarking and Self-Recovery Based on Independent Component Analysis, Lecture Notes in Computer Science, Vol. 5097, pp. 839-849.

[6] M. S. Wang and W. C. Chen, (2007), A Majority-Voting Based Watermarking Scheme for Color Image Tamper Detection and Recovery, Computer Standards & Interfaces, Vol. 29, pp. 561-571.

[7] Bas P, Chassery and Jm, Macq B, (2002), Geometrically Invariant Watermarking Using Feature Points, IEEE Trans On Image Processing, Vol. 11, No. 9, pp. 1014–28.

[8] W. Qi, X. Li, B. Yang and Daofang Cheng, (2008), Document Watermarking Scheme for Information Tracking, Journal on Communications, Vol. 29, No. 10, pp. 183–190

[9] Dawei Z, Guanrong C and Wenbo L, (2004), A Chaos-Based Robust Wavelet-Domain Watermarking Algorithm, Chaos Solitons and Fractals, Vol. 22, No. 1. pp. 47–54.

[10] G. Schirripa, C. Simonetti and L. Cozzella, (2004), Fragile Digital Watermarking by Synthetic Holograms, Proc. Of European Symposium on Optics/Fotonics In Security & Defence, London, pp. 173-182.

[11] Amir Houmansadr Et Al., (2006), Robust Content-Based Video Watermarking Exploiting Motion Entropy Masking Effect, In Proceedings of The International Conference On Signal Processing and Multimedia Applications, pp. 252-259,

[12] Kankanhalli Ms, Rajmohan, Ramakrishnan Kr, (1999), Adaptive Visible Watermarking of Images. In: IEEE International Conference on Multimedia Computing and Systems, Vol 1, pp 568–573

[13] S Radharani et. al., (2010), A Study on Watermarking Schemes for Image Authentication, International Journal of Computer Applications (0975 – 8887), Vol. 2, No.4, pp. 24-32

[14] Kay, S. and Izquierdo, E, (2001), Robust Content Based Image Watermarking, Proc. Workshop on Image Analysis for Multimedia Interactive Services.

[15] Kim, M. And Lee, W., (2004), A Content-Based Fragile

Watermarking Scheme for Image Authentication, Lecture Notes in Computer Science, Content Computing, Springer Berlin / Heidelberg, Vol. 0302/2004, pp. 258-265.

[16] Habib, M., Sarhan, S. And Rajab, L., (2005), A Robust Fragile Dual Watermarking System in The DCT Domain, Lecture Notes in Computer Science, Knowledge-Based Intelligent Information and Engineering Systems, Springer Berlin / Heidelberg, Vol. 3682/2005, pp. 548-553

[17] Chin-Shiuh Shieh et al.., (2004), Genetic Watermarking Based on Transform-Domain Techniques, Journal of Pattern Recognition, Vol. 37, pp. 555 – 565

[18] D.E. Goldberg, (1992), Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, Reading Ma.

[19] Sameh Oueslati, Adnane Cherif, and Bassel Solaiman, (2013), Multiple Binary Images Watermarking in Spatial and Frequency Domains, International Journal of Computer Theory and Engineering, Vol. 5, No. 4. pp. 598-602

[20] Zhe-Ming Lu, Dian-Guo Xu and Sheng-He Sun, (2005), Multipurpose Image Watermarking Algorithm Based on Multistage Vector Quantization, In IEEE Transactions on Image Processing, Vol. 14, No. 6, pp. 822-831. Doi: 10.1109/Tip.2005.847324

[21] Nicholas Paul Sheppard et.al., (2001), On multiple watermarking, MM&Sec '01 Proceedings of the 2001 workshop on Multimedia and security: new challenges, pp. 3-6

[22] W. S. Zhang X., (2007), Watermarking Scheme Capable of Resisting Sensitivity attack, IEEE Signal Processing Letters, Vol. 14, No. 2, pp. 125 -128.

[23] S Radharani et.al., (2011), Multiple Watermarking Scheme for Image Authentication, and Copyright Protection using Wavelet based Texture Properties and Visual Cryptography, International Journal of Computer Applications (0975 – 8887), Volume 23, No.3, pp. 29-36

[24] Houmansadr, A., Ghaemmaghami, S.: (2005), A digital image watermarking scheme based on visual cryptography, International Symposium on Telecommunications, pp. 1-5

[25] Chetan K R and S Nirmala, (2016), A novel intelligent multiple watermarking schemes for the protection of the information content of a document image, 2nd Workshop on Computer Vision Applications, Indian Institute of Technology, Guwahati

[26] Chetan K. R and S. Nirmala, (2015), An Efficient And Secure Robust Watermarking Scheme For Document Images Using Integer Wavelets And Block Coding Of Binary Watermarks. J. Inf. Sec. Appl. 24-25: pp. 13-24

[27] Chetan K R and S Nirmala, (2016), A new curvelet based blind semi-fragile watermarking scheme for authentication and tamper detection of digital images, Third International Conference on Information System Design and Intelligent Applications (INDIA 2016), Vishakapatnam.

[28] Chetan K R, S. Nirmala, A Novel Fragile Watermarking Scheme Based on Contourlets For Effective

Tamper Detection, Localization and Recovery of Handwritten Document Images, International Journal of Multidimensional Systems and Signal Processing (Communicated)

[29] Basilio et.al., (2011), Explicit Image Detection Using YCbCr Space Color Model as Skin Detection, Proceedings of the 2011 American Conference on Applied Mathematics and the 5th WSEAS International Conference on Computer Engineering and Applications, pp. 123-128

[30] Raman Maini & Dr. Himanshu Aggarwal, (2009), Study and Comparison of Various Image Edge Detection Techniques, International Journal of Image Processing (IJIP), Vol. 3, No. 1, pp. 1-12.

[31] Gadkari D., (2000), "Image Quality Analysis Using GLCM", A thesis submitted in partial fulfilment of the requirements for degree of Master of Science in Modeling and Simulation in the College Arts and Sciences at the University of Central Florida.

[32] Ming-Wei Lin, Jules-Raymond Tapamo, Baird Ndovie. (2007), A Texture-based Method for Document Segmentation and Classification. Revue Africaine de la Recherche en Informatique et Mathématiques Appliquées, INRIA, Vol. 6, pp.49-56

[33] Er. Deepak Aggarwal, (2010), An Efficient Watermarking Algorithm to improve payload and robustness without affecting Image Perceptual Quality, Journal of Computing, Volume 2, Issue 4, ISSN 2151-9617

[34] Xinshan Zhu et al., (2014), Normalized Correlation-Based Quantization Modulation for Robust Watermarking, IEEE Transactions on Multimedia, Vol. 16, No. 7, pp. 1888-1904

