

Quantization-Unaware Double JPEG Compression Detection

Ali Taimori · Farbod Razzazi · Alireza Behrad · Ali Ahmadi · Massoud
Babaie-Zadeh

Received: date / Accepted: date

Abstract The current double JPEG compression detection techniques identify whether or not an JPEG image file has undergone the compression twice, by knowing its embedded quantization table. This paper addresses another forensic scenario in which the quantization table of a JPEG file is not explicitly or reliably known, which may compel the forensic analyst to blindly reveal the recompression clues. To do this, we first statistically analyze the theory behind quantized Alternating Current (AC) modes in JPEG compression and show that the number of quantized AC modes required to detect double compression is a function of both the image's block texture and the compression's quality level in a fresh formulation. Consequently, a new double compression detection algorithm is proposed that exploits footprints introduced by all non-zero and zero AC modes based on Benford's law in a low-dimensional representation via PCA. Then, some evaluation frameworks are constructed to assess the robustness and generalization of the proposed method on

various textured images belonging to three standard databases as well as different compression quality level settings. The average F_1 -measure score on all tested databases in the proposed method is about 74% much better than the state-of-the-art performance of 67.7%. The proposed algorithm is also applicable to detect double compression from a JPEG file and localize tampered regions in actual image forgery scenarios. An implementation of our algorithms and used databases are available upon request to fellow researchers.

Keywords Double JPEG compression · image forgery detection · quality level · quantized AC modes · sparse signal · texture.

1 Introduction

In digital forensics community, the identification of image recompression history can be regarded as a clue to discover image manipulations in many forgeries [15, 37]. Among developed industrialized image compression approaches, most of digital cameras today deliver images in Joint Photographic Experts Group (JPEG) compression standard directly [42]. Hence, to manipulate an image, a photographic forger may tamper a JPEG image in one photo-editing software package and compress it again in the same well-known default format, which results in a double compressed image. Due to information loss in the lossy JPEG compression technique, double JPEG compression introduces some statistical artifacts which may be exploited in a forensic analyzer system to be able to discriminate an original single compressed JPEG image from a double compressed one or localize tampered regions in a forged JPEG image. This paper focuses on these two problems.

Ali Taimori · Farbod Razzazi
Department of Electrical and Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran 14778-93855, Iran
E-mail: a.taimori@srbiau.ac.ir; razzazi@srbiau.ac.ir

Alireza Behrad
Faculty of Engineering, Shahed University, Tehran 18651-33191, Iran
E-mail: behrad@shahed.ac.ir

Ali Ahmadi
Department of Electrical and Computer Engineering, K. N. Toosi University of Technology, Tehran 14317-14191, Iran
E-mail: ahmadi@eetd.kntu.ac.ir

Massoud Babaie-Zadeh
Department of Electrical Engineering, Sharif University of Technology, Tehran 14588-89694, Iran
E-mail: mbzadeh@sharif.edu

1.1 Related Work

In order to predict the condition whether a query JPEG image is double compressed or not, different methods have been developed [7, 10, 11, 19, 25, 28, 31, 32, 36, 38]. The ideas behind double compression detection techniques can potentially facilitate the localization of altered regions in images [8, 14, 26, 27, 43]. We categorize double compression detection methods into three groups according to footprints that are exploited for feature extraction purposes. In the first category, some of the studies are based on the assumption of non-aligned blocking artifacts during double compression creation [7, 8]. As the second category, due to the nature of non-overlapping block-based processing in JPEG encoder, there are double compression detection methods that leverage inconsistencies introduced in the boundaries of image's blocks or the lack of correlation between neighboring blocks [10, 11, 28]. In the third category, the statistical distributions of quantized Discrete Cosine Transform (DCT) coefficients in the compressed domain are investigated [10, 11, 25, 32, 32, 38].

Because of proper matching between Benford [6] and 2-D DCT coefficients distributions [35, 41], some double compression detection techniques have been analyzed the cue of Quantized DCT (QDCT) coefficients statistics of doubly compressed JPEG images [11, 25, 32]. In [25], Li et al. proposed a feature extraction approach based on the first digit attributes of individual Alternating Current (AC) modes from quantized 2-D DCT coefficients in a zig-zag order. In this seminal work, they trained a Linear Discriminant Analysis (LDA) classifier to distinguish double compression from single compression. In another study, Milani et al. suggested a learning-based multiple compression detection method [32] which is able to identify recompression stages up to four times. To do this, they analyzed the effect of decimal digits on the Most Significant Digit (MSD) distribution from low-frequency quantized AC coefficients in order to improve the accuracy of [25]. The employed coefficients of the mentioned study had been already proposed in [36] for steganography purposes. In [38], Popescu and Farid detected double compressed images via a specific peak pattern which is obtained from the magnitude of Discrete Fourier Transform (DFT) of QDCT modes. To localize tampered regions in the presence of double compression, X. H. Li et al. proposed a sliding-window image forgery detector [26] for improving the performance of the saliency map-based tampering detector of [27]. Contrary to the approaches like [25, 32, 38] which are exploited the first order statistics, by using a Markov transition probability matrix, two methods [10, 11] utilized the second-order statis-

tics of QDCT coefficients to enhance the performance of double compression detection algorithms by leveraging the inter-block correlation of coefficients.

In [19], Huang et al. developed an approach to merely deal with double compression detection in the special case that both primary and secondary Quantization Tables (QTs) are the same. Their algorithm is based on random perturbation strategy and demonstrates promising results especially in high identical compression's quality levels. In [23], Lai and Böhme revealed a characteristic called the block convergence property in repeated JPEG compressed images with the highest possible compression's quality level, denoted by $l = 100$ in [22, 30], which can be employed to solely estimate the number of recompression stages for such a setting. Therefore, the methods [19, 23] are limited to a special case of the double compression detection problem.

1.2 Motivating Forensic Scenario

In the classic forensic scenario of double JPEG compression detection, if the image under investigation is given in JPEG format, image forensics analyzer is aware of the last QT. It is achievable by reading the embedded encoding/decoding information in JPEG header file. Provided that the utilized methodology in JPEG compression scheme is explicitly known, the secondary quality level, l_2 , is also uniquely identifiable. In such a scenario for which analyzer is aware of the secondary quality level, but unaware of the primary quality level, l_1 , several semi-aware techniques have been developed to distinguish single compressed images from double compressed ones [10, 11, 25, 31].

This paper takes another case into account in which both the primary and secondary quality levels are not explicitly or reliably known and may compel the forensic analyst to blindly reveal recompression clues. We call the problem in such a case as the "quantization-unaware double JPEG compression detection". As a real-world application, in some forensic investigation scenarios, only a bitmap (BMP) image version of a previously double compressed JPEG image may be available. In this scenario, the forger may deliberately mislead the forensic analyst about the forensically useful information of compressed file such as the embedded QT and/or may provide the image on digital media for further manipulation in a photo-editing software tool like Adobe Camera Raw to, for instance, a photojournalist applies raw image processing techniques on the image. It is also likely to the QT of a JPEG metadata be noisy in a communication channel or be missing of the JPEG header file [40], so that is not reliably identifiable.

In addition to the above practical use-cases of the quantization-unaware identification, the major advantageous characteristics of quantization-unaware double JPEG compression detection are two-fold. First, the semi-aware learning-based methods of [10, 11, 25, 31] require an individual trained model for each l_2 to attain high classification performance. A motivating example is to design a real-life double JPEG compression detection system to consider, for the standard JPEG scheme [22, 30], all possible compression's quality levels in the set $\mathcal{Q} := \{l \mid l \in \mathbb{N}_1, 1 \leq l \leq 100\}$, where the symbol \mathbb{N}_1 denotes the set of all positive integer numbers. It is needed a long time, in addition to a great memory requirement, to train 100 individual classifiers which are only responsible for one JPEG encoder scheme. Instead, in the quantization-unaware case, only one global learned model is exploited. Second, the idea behind quantization-unaware double compression identification corresponds approximately to the scenario in which the image under investigation is available in spatial domain, e.g. the decoded double compressed JPEG image that has been resaved in an uncompressed form like the bitmap file. In this case, the secondary compression history is unavailable. To be aware of quantization history, one way is to design a preprocessor stage for estimating the secondary quality level and then behave with the problem as a quantization-semi-aware case. However, a new challenge arises in such a methodology, i.e. the estimation error of the secondary quality level, which its effect will propagate in a forensic analyzer system. To the best of the authors' knowledge, the available identifiers only act successfully over a previously single JPEG compressed bitmap image [12, 24, 30, 33]. In another solution, the forensic examiner may blindly identify the bitmap recompression history. Due to unawareness of the secondary quality level and/or corresponding QT, the complexity of the problem to distinguish single compression from double compression will increase. Actually, in this blind approach, the identification of recompression history is more sophisticated than the semi-blind case. Unfortunately, from the feature extraction perspective, the classification performance and generality of previously published semi-aware methods will be degraded in such a scenario, especially in the high-quality or well-textured compressed images.

1.3 Paper Contributions

In order to find a proper global model for the quantization unaware identification, we statistically analyze the basic theory behind quantized AC modes in JPEG compression scheme by using a fresh formulation. To

the best of our knowledge, none of the previous studies have analyzed the effect of the image texture on the behavior of quantized AC modes in the JPEG compression standard and have gained it to detect double compression. Only, the work [28, 31] referred to the effect of image content on the double compression detection. In this paper, we rationalize the underlying statistical behavior of the quantized AC modes to be a basis for detecting the JPEG recompression. In order to grasp an optimal performance, we show that the number of quantized AC modes required to detect double compression is related to both the image's block texture and the compression's quality level. That is, based on the amount of the last quality level and image texture complexity, especially in the high-quality or well-textured compressed images, we demonstrate that the QDCT matrix is not highly sparse, and consequently almost all of the attributes are informative. This result is what has been neglected by the related work. The approaches [11, 25, 32, 36, 38] employed only low-frequency coefficients in their feature extraction mechanisms. However, as will be mathematically shown in this paper, considering the low-frequency AC modes in the process of a given feature extraction approach is caused to miss discriminative information of single compression from double compression. If only the texture or compression's quality level of the image under investigation is low enough, this hypothesis may be true. But, the texture of image blocks has a random nature and may be different from one block to another. And, forgers usually target images with the high compression's quality levels in cyberspace to create forged images [24].

Due to these facts, we suggest a Double Compression Detection (DCD) algorithm that leverage the footprints introduced by all AC frequency modes to achieve an outstanding classification performance. Our low dimensional features are extracted by using Benford's law [6, 41] and Principal Component Analysis (PCA) [21]. We also represent the application of the proposed algorithm for detecting double compression in the quantization semi-aware scenario and localizing forged regions in manipulated JPEG images. For reproducibility of the results, an implementation of our algorithms and used databases are available upon request to fellow researchers.

The remainder of this paper is organized as follows. Section 2 statistically investigates the behavior of quantized AC modes in the JPEG compression standard. In Section 3, we analytically study the process of double compression. Starting from these analyses, Section 4 describes the proposed double compression detection algorithm. In Section 5, we empirically assess the suggested method. Finally, in Section 6, the paper is concluded.

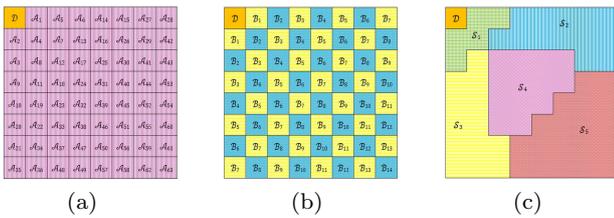


Fig. 1: Three various partitions of the mode space: (a) a partitioning based on DC (the subset \mathcal{D}) and AC modes (the subsets \mathcal{A}_1 to \mathcal{A}_{63}) in a zig-zag order, (b) a partitioning based on the notion of frequency bands of \mathcal{B}_1 to \mathcal{B}_{14} , and (c) a partitioning based on the definition presented in [17], which subsets \mathcal{S}_1 to \mathcal{S}_5 respectively correspond to low, horizontal and vertical frequencies, diagonal textures, and high frequencies information.

2 Analysis of Quantized AC Modes Behavior

To statistically analyze the characteristics of QDCT coefficients, the basic elements of the JPEG coder are investigated [19, 30] focusing on those that have forensically telltale footprints, i.e. the transform and quantization stages. For describing the characteristics of these parts, we present the definition of frequency bands for partitioning the mode space illustrated in Fig. 1 (b).

Definition 1 (frequency bands) Let the matrix $\mathbf{F} = [f_{ij}]_{b \times b} := [\mathbf{f}_1, \dots, \mathbf{f}_b]$, $\forall f_{ij} \in \{0, 1, \dots, 2(b-1)\}$, where $\mathbf{f}_{n+1} = \mathbf{f}_n + \mathbf{1}$ ($1 \leq n < b$) and $\mathbf{f}_1 = [0, 1, \dots, b-1]^T$ (chosen psycho-visually as $b = 8$). The vector $\mathbf{1}$ is the vector of all ones. The set $\mathcal{B}_k := \{(i, j) | f_{ij} = k, 1 \leq i, j \leq b\}$, $0 \leq k \leq 2(b-1)$, is defined as the elements of the k^{th} frequency band. In fact, each frequency band index is constructed from one of all possible anti-diagonals of \mathbf{F} . The mode space set, \mathcal{M} , contains all frequency bands.

Fig. 1 (a) also redefines the Direct Current (DC) and AC modes in a zig-zag order employed in the baseline JPEG coder [42] as another partition of the mode space ($\mathcal{M} = \mathcal{D} \cup \mathcal{A}$). The DC mode, \mathcal{D} , includes only a coefficient in the frequency $f_{1,1}$ and AC modes, \mathcal{A} , consist of all the others, i.e. $|\mathcal{D}| = 1$ and $|\mathcal{A}| = b^2 - 1$.

2.1 Block-Wise DCT Characteristics

Let the matrices $\mathbf{T} = [t_{ij}] \in \mathbb{R}^{b \times b}$ and $\mathbf{R}_p = [r_{ij}^p]_{b \times b}$ are the transformation kernel of DCT and the p^{th} image block of an image, respectively. The matrix $\mathbf{R}_p^c = [r_{ij}^{c,p}] \in \mathbb{R}^{b \times b}$ represents the p^{th} block in the frequency domain as $\mathbf{R}_p^c = \mathcal{F}_c \{\mathbf{R}_p\} = \mathbf{TR}_p\mathbf{T}^T$, which the linear operator \mathcal{F}_c denotes the direct Fourier-related 2-D DCT. As a characteristic of \mathbf{R}_p^c , the mean of absolute

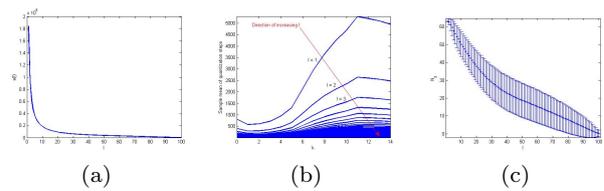


Fig. 2: (a), (b) Quantization Table (QT) characteristics: (a) the integral of QT vs. the quality level, (b) the sample mean of quantization steps vs. frequency bands for all quality levels, and (c) the number of zero entries of the QDCT matrix vs. the quality level for 200000 uniformly Gaussian distributed synthetic image blocks.

values of spatial frequencies for all bands has almost a decreasing envelope. Also, the basis vectors of \mathbf{T} are constant and independent of statistical properties of the image block. Consequently, the energy compaction property of DCT to place the signal's energy into the low frequency bands and result in a sparse signal is not necessarily optimal. And, the zig-zag sort of AC modes does not represent the optimal orientation of their sparsity. As evidence of this matter, see Fig. 1 (c) presented in [17] which groups similar frequency contents.

2.2 Quantization Table Characteristics

Let the matrix $\mathbf{Q}_l = [q_{ij}^l]_{b \times b}$, $\forall q_{ij}^l \in \mathbb{N}_1, l \in \mathcal{Q}$, is the QT corresponding to the l^{th} quality level. To characterize the behavior of the QTs proposed by Independent JPEG Group (IJG), consider the function $a(l)$ as the integral of \mathbf{Q}_l . As an characteristic of the luminance QT, the function $a(l)$ is certainly a decreasing curve, which Fig. 2 (a) plots it. As another attribute, the sample mean of quantization steps (i.e. the entries of \mathbf{Q}_l) of each frequency band has a specific curve, where, for all l , Fig. 2 (b) shows its curves. And again, all of the higher-quality level curve places below the lower-quality level one. There is an almost ascending behavior in the curves versus the frequency band number, although some exceptions exist in the starting and ending bands.

2.3 Quantized DCT Coefficients Characteristics

The QDCT matrix $\tilde{\mathbf{R}}_p = [\tilde{r}_{ij}^p] \in \mathbb{Z}^{b \times b}$ is determined as $\tilde{\mathbf{R}}_p = [\mathbf{R}_p^c \oslash \mathbf{Q}_l]$, where the notations \oslash and $[\cdot]$ stand for the entry-by-entry right division and the nearest integer, respectively. On the low-pass natural scenes such as a clear sky image, the matrix $\tilde{\mathbf{R}}_p$ is generally sparse. Almost, the amount of sparsity is inversely proportional to both l and the image texture. To show the first phenomenon, i.e. the relation between the information of

$\tilde{\mathbf{R}}_p$ and l , let $N_z = b^2 - \|\tilde{\mathbf{R}}_p\|_0$ denotes the number of zero entries of $\tilde{\mathbf{R}}_p$, which the l_0 -norm, $\|\cdot\|_0$, represents the number of non-zero elements of a vector/matrix. If the quality level in JPEG encoder increases then its corresponding quantization steps will decrease (See Figs. 2 (a) and (b)). In this case, the number of non-zero coefficients will increase, and consequently, N_z will decrease. This characteristic implies that $N_z \approx N_z(l)$, which Fig. 2 (c) empirically illustrates such a functionality for 200000 uniformly Gaussian distributed synthetic image blocks. For all l , about 99.7% of the zero entries fall in the plotted intervals. The variations of N_z at the middle quality levels are larger than other levels.

In the second phenomenon, increasing the texture is behaviorally proportional to an increase in l . That is, if the coarseness of an image's block texture increases, N_z will decrease. We use Shannon's joint entropy which is determined from Haralick's co-occurrence matrix [16, 20] as a quantitative metric to measure the texture of a block. Hence, the normalized local texture of the p^{th} block, \mathbf{R}_p , belonging to the λ^{th} image is calculated as

$$T_p^\lambda = \frac{1}{H_{\max}} \sum_{i=1}^B \sum_{j=1}^B \hat{p}_{ij} \log_2 \left(\frac{1}{\hat{p}_{ij}} \right), \quad (1)$$

where $T_p^\lambda \in [0, 1]$, $\forall p, \lambda$. The gray level co-occurrence matrix $\hat{\mathbf{P}} = [\hat{p}_{ij}]_{B \times B}$ is the joint Probability Mass Function (PMF) which is obtained by the quantized or scaled version of the input block, namely $\mathbf{G}_s = [g_{ij}^s]_{b \times b}$, so that $\hat{p}_{ij} = P(g_{r,c}^s = i, g_{r,c+1}^s = j)$, where $1 \leq i, j \leq B$, $1 \leq r \leq b$ and $1 \leq c \leq b-1$. We set the number of distinct gray levels in \mathbf{G}_s as $B = 8$, so that $g_{ij}^s \in [1, B]$, $1 \leq i, j \leq b$. In (1), H_{\max} denotes the maximum entropy of $\hat{\mathbf{P}}$, which $B = 8 \Rightarrow H_{\max} = 6$.

To qualitatively evaluate textural information of a database, we divide its images into three groups based on their measured texture content. To do this, let \bar{m}_T^λ , $\forall \lambda \in [1, N_R]$, denotes the sample mean of local texture over all blocks of the λ^{th} image, \mathbf{R}^λ , belonging to a given database, in which N_R represents the total number of images. So, the number of images with low, moderate and high textures of the given database are defined as $N_{LT} := |\{\mathbf{R}^\lambda \mid 0 \leq \bar{m}_T^\lambda < t_L, 1 \leq \lambda \leq N_R\}|$, $N_{MT} := |\{\mathbf{R}^\lambda \mid t_L \leq \bar{m}_T^\lambda < t_H, 1 \leq \lambda \leq N_R\}|$ and $N_{HT} := |\{\mathbf{R}^\lambda \mid t_H \leq \bar{m}_T^\lambda \leq 1, 1 \leq \lambda \leq N_R\}|$, respectively. We set the thresholds $t_L = \frac{1}{3}$ and $t_H = \frac{2}{3}$.

Now, to empirically show $N_z \approx N_z(T_p^\lambda)$, suppose the vector $\tilde{\mathbf{r}}_k$ includes the coefficients corresponding to the k^{th} frequency band. Then, the probability that the event of zero occurs in the k^{th} frequency band is

$$P(\tilde{\mathbf{r}}_k = \mathbf{0}) = 1 - \frac{\|\tilde{\mathbf{r}}_k\|_0}{|\mathcal{B}_k|}, \quad 0 \leq k \leq 2(b-1). \quad (2)$$

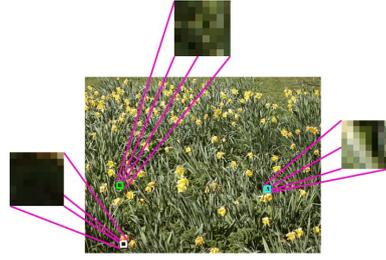


Fig. 3: Random nature of block's texture in the three magnified blocks, highlighted on the 341th image of UCID database in white, green and cyan colors which have low, moderate and high textures, respectively.

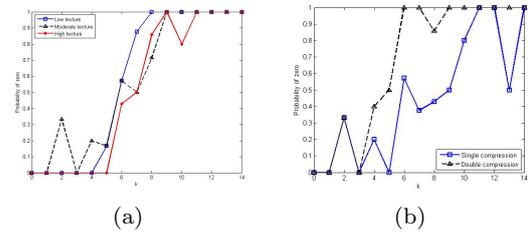


Fig. 4: (a) The probability of zero for the low, moderate and high texture blocks of Fig. 3, (b) the probabilities of zero in the image block with the moderate texture in Fig. 3 for both single and double compressions.

Example 1 (random nature of block's texture) As a real example, see Fig. 3 that represents random nature of texture in three image blocks of the 341th ensemble of Uncompressed Colour Image Database (UCID) [39], for which $T_{2891}^{341} = 0.1626$, $T_{1866}^{341} = 0.5141$ and $T_{1970}^{341} = 0.8280$, having the low ($N_z = 40$), moderate ($N_z = 37$) and high ($N_z = 33$) textures, respectively. For $l = 70$, Fig. 4 (a) plots (2) in these blocks. The sample mean \bar{m}_T^{341} is the maximum value in UCID database.

As a conclusion, depending on the amount of the last quality level and image texture, especially in the high-quality or well-textured compressed images, the matrix $\tilde{\mathbf{R}}_p$ is not necessarily sparse, i.e. $N_z = N_z(l, T_p^\lambda)$, and almost all of the frequency attributes are informative to find global features for solving the problem in the quantization-unaware case. These intuitions notify us that the required number of quantized AC modes to identify double compression is a function of l and T_p^λ .

3 Analytical Study of Double Compression

The quantization stage in the coding process of JPEG compression is the main source of lossiness [13] that propagates the quantization error, due to rounding in

frequency domain. Hence, we statistically analyzed double compression regardless of other negligible sources of error [19, 30]. Based on the basic building elements of JPEG scheme [30], for all t in \mathbb{N}_0 (in which the symbol \mathbb{N}_0 denotes the set of all nonnegative integer numbers), we recursively model the recompression process as

$$\hat{\mathbf{R}}_p^{(t+1)} = \mathcal{T}([\mathcal{F}_c^{-1}\{([\mathcal{F}_c[\hat{\mathbf{R}}_p^{(t)}] \oplus \mathbf{Q}_{l_{t+1}}]) \odot \mathbf{Q}_{l_{t+1}}\}], \quad (3)$$

where $\hat{\mathbf{R}}_p^{(t+1)}$ represents the p^{th} reconstructed image blocks in spatial domain after $(t+1)^{\text{th}}$ compression, for which $\hat{\mathbf{R}}_p^{(0)} := \mathbf{R}_p$. The symbols \odot , \mathcal{F}_c^{-1} and \mathcal{T} stand for Hadamard product, the inverse Fourier-related 2-D DCT and the truncation function, respectively. For single and double compressions in pixel domain, we have

$$\hat{\mathbf{R}}_p^{(1)} \approx \mathcal{F}_c^{-1} \left\{ \tilde{\mathbf{R}}_p \odot \mathbf{Q}_{l_1} \right\}, \text{ and}, \quad (4)$$

$$\begin{aligned} \hat{\mathbf{R}}_p^{(2)} &\approx \mathcal{F}_c^{-1} \left\{ \left[\mathcal{F}_c \left[\mathcal{F}_c^{-1} \left(\tilde{\mathbf{R}}_p \odot \mathbf{Q}_{l_1} \right) \right] \oplus \mathbf{Q}_{l_2} \right] \odot \mathbf{Q}_{l_2} \right\} \\ &\approx \mathcal{F}_c^{-1} \left\{ \left[\tilde{\mathbf{R}}_p \odot \mathbf{Q}_{l_1} \oplus \mathbf{Q}_{l_2} \right] \odot \mathbf{Q}_{l_2} \right\}. \end{aligned} \quad (5)$$

In the following lemma, we study the behavior of the quantization error as well as the effect of AC modes selection as individual distinguishable features on the detection error of double compression (positive class, C_2) as single compression (negative class, C_1) regardless of other factors in machine learning systems. Generally, from a forensic analyzer viewpoint, to identify image recompression history, the cost of the misclassification of the class C_2 as C_1 , producing the type II error or False Negative (FN), is more expensive than the class C_1 as C_2 , producing the type I error or False Positive (FP). This is because a predicted double compressed image may be investigated in further forensic tools for localizing its probable tampered regions, whereas a detected single compressed image may be known as an authentic signal. Therefore, our emphasis is on FN.

Lemma 1 (AC modes effect on detection error)

Let the set $\mathcal{M} = \{\mathcal{D}, \mathcal{A}\}$ be as the mode space in which the sets \mathcal{D} and \mathcal{A} denote the single-element DC mode and all AC modes, respectively. If the set \mathcal{A} is partitioned into n subsets so that $\mathcal{A}_i \cap \mathcal{A}_j = \emptyset$, $\forall i \neq j$, then, the misclassification error probability of double compressed images as single compressed ones satisfies

$$P_e(C_1|C_2, \mathcal{A}) \leq P_e(C_1|C_2, \mathcal{A}_k), \forall k = 1, \dots, n. \quad (6)$$

Proof In (5), we first define the matrix $\Phi := \mathbf{Q}_{l_1} \oplus \mathbf{Q}_{l_2}$, $\forall \phi_{ij} \in \mathbb{R}^+$. It can be observed that $\mathcal{A} = \bigcup_{k=1}^n \mathcal{A}_k$, hence,

$$P(\langle \Phi \rangle \neq \mathbf{O} | \mathcal{A}) = \sum_{k=1}^n P(\langle \Phi \rangle \neq \mathbf{O} | \mathcal{A}_k), \quad (7)$$

where the matrix \mathbf{O} is the zero matrix and $P(\langle \Phi \rangle \neq \mathbf{O} | \mathcal{A}_k) := \frac{1}{|\mathcal{A}_k|} |\{(i, j) \in \mathcal{A}_k | \langle \phi_{ij} \rangle \neq 0\}|$, $\forall k$. The symbols $\langle \cdot \rangle$ and $[\cdot]$ denote the fractional part and the floor function, respectively. To prove, three possible cases exist for l_2 in a doubly compressed JPEG image as follows.

- Case I: If $l_1 = l_2$, then, $q_{ij}^{l_1} = q_{ij}^{l_2}$, $\forall i, j$, and $\Phi = \mathbf{J}$, which the matrix \mathbf{J} is the all-ones matrix. We conclude $\hat{\mathbf{R}}_p^{(2)} = \hat{\mathbf{R}}_p^{(1)}$. Hence, for all k , $P(\langle \Phi \rangle \neq \mathbf{O} | \mathcal{A}) = P(\langle \Phi \rangle \neq \mathbf{O} | \mathcal{A}_k) = 0$ and all doubly compressed images classify as singly compressed ones ($P_e(C_1|C_2, \mathcal{A}) = P_e(C_1|C_2, \mathcal{A}_k) = 1$). As a result, because the trace of quantization error does not exist in the second compression, there is no meaningful discriminative information for learning to classify the class C_1 from C_2 correctly. So, for a double compressed block, the number of zeros in $\tilde{\mathbf{R}}_p$ is equal to its single compressed counterpart.

- Case II: If $l_1 < l_2$, then, $q_{ij}^{l_1} \geq q_{ij}^{l_2}$ and $1 \leq \phi_{ij} < \infty$, $\forall i, j$. Now, consider the worst situation occurs, i.e. all quantization steps in the primary compression are integer multiples of the secondary compression counterpart. In other words, $\exists \alpha \in \mathbb{N}_1$ so that $\mathbf{Q}_{l_1} = \alpha \mathbf{Q}_{l_2}$. By substituting \mathbf{Q}_{l_1} in Φ , this situation yields $\Phi = \alpha \mathbf{J}$. Therefore, $\hat{\mathbf{R}}_p^{(2)} = \hat{\mathbf{R}}_p^{(1)}$ and identical results with the case I will happen. Consequently, for all k , $P(\langle \Phi \rangle \neq \mathbf{O} | \mathcal{A}) \geq P(\langle \Phi \rangle \neq \mathbf{O} | \mathcal{A}_k)$ and $P_e(C_1|C_2, \mathcal{A}) \leq P_e(C_1|C_2, \mathcal{A}_k)$. Thus, exploiting more AC modes information result in more non-integer multiples to help the learning system to alleviate the misclassification error of the class C_2 as C_1 . Also, the value N_z of a double compressed block is less than or equal to its single compressed counterpart.

- Case III: If $l_1 > l_2$, then, $q_{ij}^{l_1} \leq q_{ij}^{l_2}$ and $0 < \phi_{ij} \leq 1$, $\forall i, j$. Here again, suppose the worst circumstance occurs, i.e. all quantization steps in the secondary compression are integer multiples of the primary compression counterpart. In other words, $\exists \beta \in \mathbb{N}_1$ so that $\mathbf{Q}_{l_1} = \frac{1}{\beta} \mathbf{Q}_{l_2}$. Such a situation yields $\Phi = \frac{1}{\beta} \mathbf{J}$, and then, $\hat{\mathbf{R}}_p^{(2)} = \mathcal{F}_c^{-1} \left\{ \left[\frac{1}{\beta} \tilde{\mathbf{R}}_p \right] \odot \beta \mathbf{Q}_{l_1} \right\} \neq \hat{\mathbf{R}}_p^{(1)}$. Consequently, for all k , $P(\langle \Phi \rangle \neq \mathbf{O} | \mathcal{A}) > P(\langle \Phi \rangle \neq \mathbf{O} | \mathcal{A}_k)$ and $P_e(C_1|C_2, \mathcal{A}) < P_e(C_1|C_2, \mathcal{A}_k)$. We observe that even for the worst situation, the information loss yet exists. As a general result, for feature extraction approaches that leverage all AC frequency modes, we expect to attain a stronger improvement in the performance of the case III than other two cases. Additionally, the value N_z of a doubly compressed block is more than or equal to its singly compressed counterpart. To observe this difference, Fig. 4 (b) plots $P(\tilde{\mathbf{r}}_k = \mathbf{0})$, $\forall k$ in the block with the moderate texture of Example 1 for single compression with $l_1 = 80$ and double compression with $(l_1 = 80, l_2 = 55)$. These curves show the discrimination between the classes C_1 and C_2 especially in middle and high frequencies. \square

4 Double Compression Detection Algorithm

Based on the analyses of Sections 2 and 3 as well as the following intuitions, we have defined the feature space in the proposed DCD algorithm.

a) JPEG compression performs the identical process on dense and sparse image blocks, and does not consider their frequency contents. On the other hand, different camera manufacturers as well as image editing software packages do not define the QT by the standard [1]. Therefore, it is required to attributes of double compression detection be as independent as possible from such variations. Independent of which definition for the QT is used, Lemma 1 denotes that attributes may be spread in different partitions of the mode space.

b) To consider statistics of the non-zero QDCT coefficients, we firstly suggest to exploit the effect of all single-frequency modes to grasp spread attributes. Then, a compact representation of these attributes are presented by Benford's law. To this intent, we determine the FSD distribution coefficients of the individual quantized AC mode as the first set of features to detect double compression. These features cope with dense blocks and reduce dependency to the employed QT.

c) In the sparse blocks, the zero entries of QDCT coefficients for a double compressed image block may be different from a single compressed version. To take into account the number of zero QDCT coefficients, N_z , we employ the Second Significant Digit (SSD) to be able to capture discriminative information laid in the zero coefficients as the second set of features. This work has the benefit that not only the zero entries of QDCT coefficients are counted, but also the zero values in the SSD of non-zero coefficients are considered.

4.1 Extracting Features

By the above explanations, to extract features, at first, we construct the quantized AC modes as individual vectors. If the vector $\mathbf{m}_a^n = [m_1^{n,a}, \dots, m_{N_p}^{n,a}]^T$ integrates the a^{th} quantized AC mode, for all blocks of the n^{th} image of a given database, then

$$\mathbf{m}_a^n = \{\tilde{r}_{ij}^p | (i, j) \in \mathcal{A}, i = i_a, j = j_a\}_{p=1}^{N_p}, 1 \leq a < |\mathcal{A}|.$$

In the next step, we estimate the real distributions of the FSD and the digit zero from the SSD based on Benford's law. For this purpose, let the function $D_L(X)$ represents the L^{th} significant decimal digit of the random variable $X \in \mathbb{R}$, then $D_L(X) \sim B_L$, where B_L is the L^{th} digit Benford distribution. Benford's law de-

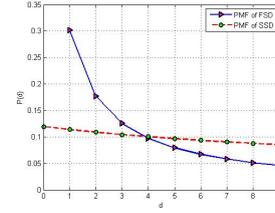


Fig. 5: The probability curves of the first and the second significant digits.

scribes the behavior of PMF as

$$P(D_L(X) = d) = \begin{cases} \log_{10} \left(1 + \frac{1}{d}\right), & L = 1 \wedge d \in \mathbb{D}_1 \\ \sum_{k=10^{L-2}}^{10^{L-1}-1} \log_{10} \left(1 + \frac{1}{10k+d}\right), & L > 1 \wedge d \in \mathbb{D}_0 \end{cases} \quad (8)$$

for which $\mathbb{D}_0 := \{0, 1, \dots, 9\}$ and $\mathbb{D}_1 := \{1, 2, \dots, 9\}$. The significant decimal digit function can be explicitly calculated [6] as follows

$$D_L(X) := \begin{cases} 0, & X = 0, L > 1 \\ \lfloor 10^{L-1} S(X) \rfloor - 10 \lfloor 10^{L-2} S(X) \rfloor, & X \neq 0, \forall L \end{cases} \quad (9)$$

where $S(X)$ represents the decimal significand function

$$S(X) := \begin{cases} 0, & X = 0 \\ 10^{\log_{10}|X| - \lfloor \log_{10}|X| \rfloor}, & X \neq 0 \end{cases} \quad (10)$$

Please note that based on the definition presented in (9), $D_1(X)$, is only counted for the non-zero real number X . But, in $D_2(X)$, both SSD of zero and non-zero numbers are considered, so that, we take into account $D_2(0) := 0$ for convenience [6]. For example, $D_2(0) = 0, D_2(105) = 0, D_2(0.000702) = 0, D_2(68) = 8, D_2(-94) = 4$. Fig. 5 plots probability curves of the first and the second significant digits.

Benford's law has four basic properties that help us to find whether significant digits of the random variable X follow Benford distribution [6, 41]. These properties are known as uniform distribution, scale-invariance, base-invariance and sum-invariance characterizations. Here, we define $m_i^{n,a} := X$, where $m_i^{n,a} \in \mathbb{Z}, \forall i$ and utilize the histogram estimator to determine the empirical PMF of the vector \mathbf{m}_a^n as

$$\hat{\mathbf{P}}_{\mathbf{m}_a^n} = [\hat{p}_1^{\mathbf{m}_a^n}, \dots, \hat{p}_{N_d-1}^{\mathbf{m}_a^n}, \hat{p}_{N_d}^{\mathbf{m}_a^n}]^T \in \mathbb{R}^{N_d}, 1 \leq a \leq |\mathcal{A}|,$$

in which the $\hat{p}_1^{\mathbf{m}_a^n}$ to $\hat{p}_{N_d-1}^{\mathbf{m}_a^n}$ entries denote the probability distribution of the FSD, the $\hat{p}_{N_d}^{\mathbf{m}_a^n}$ entry is the probability distribution of the digit zero from the SSD and the variable $N_d = N_d^F + N_d^S$ represents the total number of utilized significant digits in the proposed feature extraction method. The parameters N_d^F and N_d^S are the

number of digits employed from the FSD and the SSD, respectively, which $N_d^F = 9$ and $N_d^S = 1$. Ultimately, the feature vector of the n^{th} image is constructed as $\mathbf{x}_n = \text{vec}([\hat{\mathbf{p}}_{\mathbf{m}_1^n}, \dots, \hat{\mathbf{p}}_{\mathbf{m}_{|\mathcal{A}|}^n}]) \in \mathbb{R}^K$, where the function $\text{vec}(\cdot)$ denotes the vector operator and $K = N_d \times |\mathcal{A}|$ is the initial feature dimensions, i.e. $K = 630$.

4.2 PCA Dimensionality Reduction

As a benefit of our initial features, it exploit existing information into total space of AC modes. However, its dimensions may cause the curse of dimensionality, especially in small data (e.g. for forgery localization purposes). To solve this problem, it is needed to suppress redundant information of sparse blocks by an appropriate dimensionality reduction technique. Hence, in the proposed method, we suggest the venerable data-driven PCA which is a variant of the optimal KLT transform [21], to appropriately encode information of different frequency bands into a lower-dimensional representation. If the matrix $\mathbf{V} \in \mathbb{R}^{K \times d_r}$ represents the PCA transformation kernel of a given model, then the n^{th} training/testing dimensionality reduced feature is $\mathbf{y}_n = \mathbf{V}^T \mathbf{x}_n$, so that $d_r \leq K$. We calculated the mapping models of the unsupervised PCA dimensionality reduction technique from learning instances without their labels. After applying PCA, the dimensionality reduced features are renormalized in the range $[0, 1]$ by dividing DC-free features to their standard deviations, i.e. for the n^{th} training/testing low-dimensional sample, \mathbf{y}_n ,

$$\tilde{\mathbf{y}}_n = (\mathbf{y}_n - \mu) \oplus \sigma, \quad (11)$$

in which μ and σ denote the sample mean and standard deviation vectors, respectively, determined by all learning low-dimensional ensembles.

In the proposed DCD algorithm, feature vectors of single and double compressed images have fed to a classifier to determine the learned model. To do this, various classifiers may be used. Here, the baseline linear Support Vector Machines (SVMs) model is utilized. Considering both quantization- unaware and semi-aware scenarios, the learning phase of the proposed DCD algorithm is summarized in Algorithm 1. Fig. 6 also shows the block diagram of the suggested DCD approach in the testing phase.

4.3 Image Forgery Detection

By using the proposed method, we can localize tampered regions in manipulated JPEG images. To do this, at first, we scan the whole blocks of the questionable

Algorithm 1 The learning process of the proposed double JPEG compression detection algorithm

- 1: **Input:** Learning images set of a given database.
- 2: Dump QDCT coefficients of an image.
- 3: Integrate each quantized AC mode pertaining to all blocks of an image into an individual vector.
- 4: Estimate PMF of the FSD (digits 1 to 9) for each vector.
- 5: Estimate PMF of the digit 0 from the SSD for each vector.
- 6: Create feature vector by unifying distributions obtained from Steps 4 and 5.
- 7: Apply PCA for the low-dimensional model and then, renormalize the dimensionally reduced features.
- 8: Feed feature vectors of all training images to a binary SVMs classifier to determine the learned model.
- 9: **Output:** The learned model.

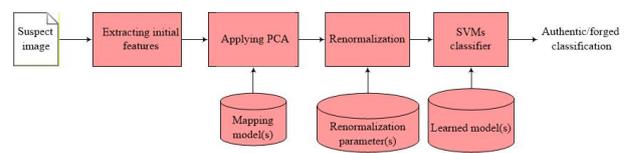


Fig. 6: The block diagram of the proposed DCD approach in the testing phase.

Algorithm 2 The pseudo-code of the proposed image forgery detection algorithm

- 1: **Input:** The questionable image \mathbf{I} in JPEG format.
- 2: Read coder information and then, dump the QDCT coefficients matrix, \mathbf{Q}_l , of the image \mathbf{I} .
- 3: **for** $p \leftarrow 1, N_p$ **do**
- 4: Put the center of a $\omega \times \omega$ mask on the p^{th} block of \mathbf{Q}_l .
- 5: Extract features.
- 6: Feed the feature vector to the SVMs classifier with the related coder quality level.
- 7: **if** \hat{o} is double compression **then**
- 8: Set $\mathbf{W}\{p\}$ with the $b \times b$ zero matrix.
- 9: **else**
- 10: Set $\mathbf{W}\{p\}$ with the $b \times b$ all-ones matrix.
- 11: **end if**
- 12: **end for**
- 13: Apply the post-processing stage to the binary image \mathbf{W} for creating the ultimate segmented image \mathbf{S} .
- 14: **Output:** The segmented image \mathbf{S} .

JPEG image \mathbf{I} , with the height of h and the width of w , by using a sliding square mask with the length of ω blocks. Afterwards, for each mask, the proposed feature vector is extracted. Then, we predict the class label of the central block, \hat{o} , via the SVMs classifier. For the image \mathbf{I} , the procedure repeats N_p times, where $N_p = \lceil \frac{h}{\omega} \rceil \times \lceil \frac{w}{\omega} \rceil$ denotes the total number of its partitions in which $\lceil \cdot \rceil$ represents the ceiling function. Ultimately, a post-processing stage is applied to connect breaks, remove noisy segments and fill holes. Algorithm 2 describes the pseudo-code of the proposed image forgery detection algorithm.

5 Experiments

Although this paper does not provide a complete mathematical proof for each part, we evaluate the proposed approach in this section with extensive experiments to support our findings. To empirically assess the robustness and generalization capabilities of the proposed double JPEG compression detection method on various textured images belonging to three standard databases as well as different compression’s quality levels, a common train/test benchmark framework is constructed. Considering both quantization- unaware and semi-aware scenarios, in this framework, we have fairly compared our DCD method to a set of competing state-of-the-art approaches. To compare the proposed feature extraction part with other related methods, we have implemented the feature extraction method of [25], the feature selection algorithm of [32], and Dong’s features [11] and then applied them to the same datasets and the baseline linear SVMs classifier for a fair comparison. To localize tampered regions, we have also compared the proposed forgery detector with the algorithm [26]. In experiments, we employed three standard real-world raw image databases in order to train and test the described classification system as follows.

- Database I: UCID database containing raw images that were captured by a Minolta Dimage 5 digital color camera from natural scenes with Tagged Image File Format (TIFF) extension [2].
- Database II: Never-compressed Color Image Database (hereafter, in the paper, called as NCID) [28, 29] was created by cropping 5000 original raw RGB images with TIFF extension, image dimensions of $\max(h, w) \times \min(h, w) = 670 \times 480$ and 8-bit per each color channel. Among cropped images, high complexity patches were intentionally selected and resaved with BMP extension and patch dimensions of 256×256 by Liu et al. [3].
- Database III: McGill Calibrated Colour Image Database (CCID) [34] encompasses color images in 9 different categories taken with two Nikon Coolpix 5700 digital cameras from natural scenes with both TIFF and JPEG extensions [4].

Fig. 7 shows representative sample images of UCID, NCID and CCID databases. In Table 1, the important statistics of the utilized various image databases are summarized. As is obvious, among databases, NCID has the most textural information and UCID has the least textural information.



Fig. 7: Representative sample images of the utilized databases: (a) UCID, (b) NCID, and (c) CCID.

Table 1: The statistics of the utilized image databases

Database	Textural information			N_R	Image dimensions
	N_{LT}	N_{MT}	N_{HT}		
UCID	1076	262	0	1338	512×384
NCID	2412	2455	283	5150	256×256
CCID	792	303	1	1096 [†]	768×576

[†] The total number of raw images in 9 categories are 1152 for which 56 images are the same. In experiments, we removed these repeated images.

5.1 Original and Altered Image Data Collection

To provide single and double compressed JPEG image patches, we firstly compressed each raw image of an individual aforementioned database with quality levels belonging to the predefined set

$$\mathcal{L}_{t+1} = \{l_{t+1} \mid l_{t+1} = l_{\min} + s_l \times k\}_{k=0}^9, \forall t \in \mathbb{N}_0.$$

In experiments, we chose the lowest quality levels, $l_{\min} = 50$, and the step size $s_l = 5$. This process itself produced $|\mathcal{L}_1|$ single compressed JPEG image databases. Then, each single compressed JPEG images database was re-compressed with the same already quality level settings that created $|\mathcal{L}_1| \times |\mathcal{L}_2|$ double compressed JPEG images databases. In other words, in this case, C_1 and C_2 classes have 10 and 100 types, respectively. Based on the holdout cross-validation approach, the number of learning data for each type of individual class is selected as $N_L = \lfloor \alpha_L \times N_R \rfloor$ where $\alpha_L \in (0, 1]$ determines the ratio of learning data utilized in the training phase. In our experiments, we set $\alpha_L = \frac{3}{4}$. Hence, the total number of learning images were $110 \times N_L$. The corresponding remainder data contained the testing set.

Fig. 8 illustrates the process of data gathering described above. Due to such a skewed class distribution, in the training phase, we encountered with a two-class imbalanced learning problem with between-class imbalance ratio 10:1. Different mechanisms may be considered to deal with this problem [18]. Here, we utilized a cost-sensitive linear SVMs classifier to grasp a robust and stable classification performance. For this purpose, we used LIBSVM toolbox [5, 9]. Therefore, in the learning process of different double compression detection

methods for the quantization-unaware scenario, we figured out the weight of classes as $w_{C_1} \approx 0.9091$ and $w_{C_2} \approx 0.0909$, with the unit cost.

5.2 Detection Efficiency on Different Quality Levels

As the first experiment, we used the well-known CCID database to train and test the aforementioned benchmark classification system. Table 2 shows the accuracy results of the proposed feature extraction method for detecting both singly compressed images, $\forall l_1 \in \mathcal{L}_1$, and doubly compressed images, $\forall l_1 \in \mathcal{L}_1, \forall l_2 \in \mathcal{L}_2$, in terms of percentage. In the table, A_{C_1} and A_{C_2} denote the accuracy for the C_1 and C_2 classes, respectively. Also, the parameters l_1 and l_2 represent ground truth labels of the first and the second compression's quality levels, respectively. Based on the discussed train and test sets, the confusion matrix is simply determinable from such a table, where in Section 5.4, we have referred to its important criteria. Tables 3, 4 and 5 tabulate the detection results of the methods [11, 25, 32] on CCID database, respectively. For ease of comparison, among the compared methods listed in Tables 2, 3, 4 and 5, the highlighted values in the bold type represent the best performance for each l_1 and (l_1, l_2) in single and double compressions, respectively. The results demonstrate the suggested features act better than the competitors, especially in the challenging case $l_1 > l_2$.

To examine the effectiveness of various methods in the high compression's quality levels, consider the sub-table for which $l_1, l_2 \in \{75, \dots, 95\}$, in Tables 2, 3, 4 and 5, i.e. the lightweight compression which are utilized frequently in JPEG images. In this situation, we determined the arithmetic mean accuracy in terms of percentage for the class C_2 to explore detection efficiency. Fig. 9 plots this criterion for all databases. The chart demonstrates that the proposed approach outperforms its competitors for all databases. It can be concluded that, to detect double compressed JPEG images with the high compression's quality levels, the proposed footprints are more robust than the features introduced in the competing methods.

5.3 Detection Efficiency on Different Textures

In order to investigate the performance of different methods on various textured images, we calculated the probability of the type II error normalized over all testing instances, $P_e(C_1|C_2) := \frac{FN}{TP+FN+FP+TN}$, for databases I, II and III as shown in Fig. 10 in terms of percentage. TP and TN stand for True Positive and True Negative, respectively. Obviously, for all databases, the proposed

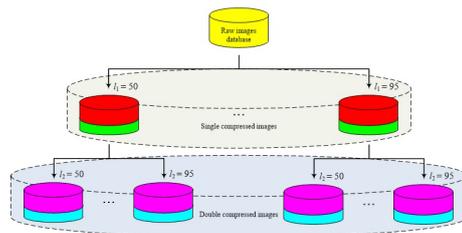


Fig. 8: The process of creating single and double compressed JPEG images.

Table 2: The accuracy results of the proposed method on CCID database in terms of percentage

$[A_{C_2}]$		l_1									
		50	55	60	65	70	75	80	85	90	95
l_2	50	7	9	3	42	11	89	96	28	30	24
	55	26	7	28	7	5	2	3	2	80	4
	60	63	9	11	45	56	16	95	54	43	64
	65	80	43	40	11	92	27	16	54	15	5
	70	81	84	85	56	11	60	16	99	88	64
	75	93	97	93	96	24	5	90	3	91	14
	80	99	96	88	88	97	76	4	83	55	67
	85	98	98	97	95	97	96	96	3	78	86
	90	96	99	99	96	96	94	98	94	1	78
	95	97	97	98	93	97	96	97	93	99	4
$[A_{C_1}]$		94	93	89	89	89	95	96	97	99	97

Table 3: The accuracy results of Li's approach [25] on CCID database in terms of percentage

$[A_{C_2}]$		l_1									
		50	55	60	65	70	75	80	85	90	95
l_2	50	9	7	3	27	6	96	97	21	18	8
	55	31	8	32	3	2	1	7	3	70	8
	60	44	18	7	85	46	9	84	60	11	18
	65	84	13	34	9	85	1	2	54	3	5
	70	91	90	95	15	6	35	0	99	74	7
	75	97	99	95	99	16	2	85	0	23	1
	80	99	97	92	95	98	74	1	79	0	3
	85	99	100	99	98	97	96	93	0	56	8
	90	97	97	99	95	98	97	100	93	0	0
	95	96	97	99	93	100	95	95	97	99	0
$[A_{C_1}]$		91	92	93	91	93	98	99	100	100	100

Table 4: The accuracy results of Milani's method [32] on CCID database in terms of percentage

$[A_{C_2}]$		l_1									
		50	55	60	65	70	75	80	85	90	95
l_2	50	20	11	13	27	85	14	84	23	24	21
	55	24	19	15	9	19	16	9	11	23	19
	60	11	42	15	29	22	45	24	55	11	16
	65	83	8	45	13	27	10	22	14	12	14
	70	5	78	58	8	11	54	5	24	13	11
	75	8	32	45	66	7	10	94	6	19	9
	80	93	84	66	15	60	0	10	46	13	11
	85	95	99	99	39	1	16	23	7	9	6
	90	26	86	78	97	10	99	3	100	9	19
	95	17	89	84	69	88	89	73	100	79	3
$[A_{C_1}]$		80	81	85	87	89	89	90	93	92	97

method is superior to the competitors. In the test on NCID database having the most textural information, for each l_1 in singly compressed images and (l_1, l_2) in doubly compressed images, the number of images with low, moderate and high textures were as $N_{LT} = 647$,

Table 5: The accuracy results of Dong’s method [11] on CCID database in terms of percentage

$ A_{C_2} $		l_1									
		50	55	60	65	70	75	80	85	90	95
l_2	50	15	14	4	40	16	95	96	22	32	12
	55	33	12	35	1	3	1	8	4	62	12
	60	50	24	12	76	49	19	82	63	17	27
	65	84	14	45	11	86	1	2	55	2	4
	70	89	88	93	23	8	38	0	98	81	8
	75	98	99	93	99	16	2	86	0	23	2
	80	99	97	92	95	99	74	1	81	0	2
	85	99	99	99	99	95	95	96	0	56	11
	90	97	98	100	96	99	97	99	94	0	0
	95	97	99	99	95	99	94	95	97	99	0
$ A_{C_1} $		85	88	88	89	92	98	99	100	100	100

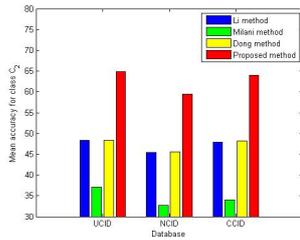
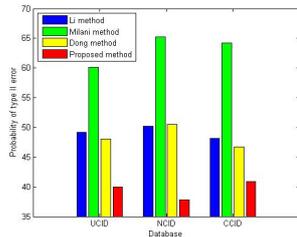
Fig. 9: The arithmetic mean accuracy for the class C_2 in terms of percentage in lightweight image compressions for different methods on databases I, II and III.

Fig. 10: The probability of the type II error in % for different methods on databases I, II and III.

$N_{MT} = 564$ and $N_{HT} = 76$, respectively. Please note that, in NCID database, the type II error has grown in the methods [25], [32] and [11], whereas for the proposed method this has decreased. Moreover, the error difference between the proposed approach with the best one of the competitors reveals that, for NCID database, this divergence rate is maximum and equal to 12.3%. These results demonstrate that the superiority of the proposed method is more important than the competitors on high textured images. In fact, those are empirical evidences of Lemma 1.

5.4 Overall Performance Measurement on Databases

Similar to the reported detailed experiment for different compression’s quality level settings on CCID database

Table 6: The overall performance metrics of different methods for quantization-unaware double compression detecting on databases I, II and III in percentage

Method	Database	Performance metrics					
		Recall			Precision	F_1 -measure	
		$l_1 < l_2$	$l_1 = l_2$	$l_1 > l_2$			Total
Li’s method [25]	UCID	85.71	3.92	26.43	50.85	99.22	67.24
	NCID	83.97	8.17	24.98	49.85	98.40	66.17
	CCID	84.54	4.23	29.78	51.87	99.18	68.12
Milani’s method [32]	UCID	55.18	17.57	29.61	39.91	95.73	56.33
	NCID	50.47	14.89	23.40	34.73	95.82	50.98
	CCID	53.24	11.61	23.64	35.76	96.87	52.23
Dong’s method [11]	UCID	86.05	5.63	28.22	51.99	98.94	68.16
	NCID	83.81	8.32	24.25	49.46	98.35	65.82
	CCID	85.28	6.17	31.70	53.26	98.86	69.23
Proposed approach	UCID	86.89	8.86	44.48	60.00	98.56	74.59
	NCID	87.03	9.94	40.46	58.37	98.34	73.25
	CCID	85.13	6.39	44.86	59.13	98.96	74.03

in Section 5.2, we also assessed the proposed and competing methods on two other databases. Here, to save space in the paper, three criteria employed as standard performance metrics, including Recall = $\frac{TP}{TP+FN}$, Precision = $\frac{TP}{TP+FP}$ and F_1 -measure = $\frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$ to appropriately investigate the performance of different methods. The higher value of each metrics means the better performance. In Table 6, the performance metrics pertaining to the proposed and competing methods are expressed as percentages for databases I, II and III. About the recall criterion which represents the performance of the double compression class, in addition to the calculation of the total recall rate, this metrics has independently been measured for cases I, II and III presented in Lemma 1. In the table, the highlighted values in the bold type indicate the best score. Experiments on three image databases demonstrated that the results of the proposed algorithm which were utilized for the problem of quantization-unaware double JPEG compression detection are promising and statistically superior, even, to its nearest competitor, i.e. Dong’s feature extraction method in almost all cases. The average F_1 -measure on three databases is about 74% in the proposed approach which is much better than the performance of 67.7% in the method [11]. Although, due to uncertainty, the F_1 -measure score of the proposed approach has degraded a bit by increasing textural information of databases, this degradation is more for the competitors, especially for NCID database. This fact shows that the proposed method is more robust against textured images, too.

5.5 Effect of Dimensionality Reduction

Because the proposed features take into account the effects of block’s texture and compression’s quality level, we experimentally demonstrate here that a low dimen-

sional representation of our features is yet more efficient than the competitors. In order to evaluate the proposed low-dimensional model of the DCD algorithm, we intentionally set the ultimate feature dimensions for PCA dimensionality reduction equal to $d_r = 200$ for comparing to the best competitor with the feature dimensions of 200 [11]. Table 7 presents the performance of the proposed low-dimensional features on CCID database in comparison with other methods. The results show that, by using PCA dimensionality reduction technique, the performance in the suggested low-dimensional model is yet more than the compared methods. It also demonstrates that dimensionality reduction after exploiting initial attributes of all single-frequency modes donates the superior performance than the competitors.

5.6 Influence of SSD-Related Features

In this experiment, we omitted the SSD-related features as the second set of attributes from the 630-element feature vector. This results in a 567-dimensional vector containing the FSD distributions of all AC frequency modes. Fig. 11 compares the performance of our 567-dimensional features against the proposed features with $K = 630$ for two different databases. Here, in addition to UCID database, we employed another database called Raw Color Image Database (RCID) including $N_R = 208$ sparse images. The authors have captured this database by a Canon Eos 550D digital color camera from natural scenes with TIFF extension and full capacity of resolution, i.e. the dimensions 5184×3456 or vice versa¹. For RCID database, $N_{LT} = 205$, $N_{MT} = 3$ and $N_{HT} = 0$, which is sparser than UCID database.

The results show the features with $K = 630$ slightly outperform the first set of attributes. The improvement values on UCID and RCID databases are 0.11% and 0.15%, respectively. As a result, the features pertaining to the digit 0 from SSD for sparse images with ample zeros in the frequency domain create better discrimination in double compression detection. The little performance degradation of our 567-element feature vector in comparison to all attributes is due to the fact that the number of zero coefficients for the class C_2 differs among cases I, II and III in Lemma 1; whereas our considered global model does not gain such intra-class variations and groups all kinds in one class. It seems that taking into account these kinds of the class C_2 as independently as possible will further strengthen the discrimination role of the SSD-related features.

¹ RCID database is publicly available to fellow academic researchers. To access it, please contact behrad@shahed.ac.ir.

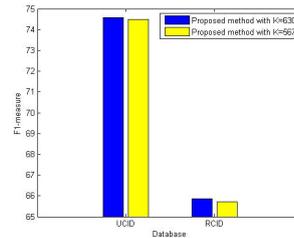


Fig. 11: The performance of the proposed features with ($K = 630$) and without ($K = 567$) SSD-related attributes in terms of percentage for two databases.

5.7 Test in Quantization-Semi-Aware Scenario

In the aforementioned experiments, we completely evaluated different double compression detection algorithms in the quantization-unaware forensic scenario. In this experiment, we have also assessed the effectiveness of the proposed DCD algorithm in the quantization-semi-aware scenario. It is important to note that, for each l_2 , we determined an individual learned model. Table 8 reports the performance of different methods on UCID database. Generally speaking, the overall performance on the quantization-semi-aware case is better than the quantization-unaware scenario. Obviously, the performance of the proposed algorithm has improved on the methods [11, 25, 32] in the quantization-semi-aware double compression detection scenario, too.

5.8 Image Forgery Detection Tests

We also examined the application of the proposed algorithm to actual forensic investigations. In this experiment, we firstly trained the individual dimensionality reduced models by all possible single and double compressed images of NCID database. Then, we provided a test-set of forged JPEG images via CCID database, each of which was consisted of both authentic and tampered regions to evaluate the performance. For instance, Figs. 12 (a), (b) and (c) portray six original JPEG images, their forged versions and detection results of tampered regions, respectively. The figure shows that the results of forgery detection are promising. In experiments, we set the parameters as $\omega = 9$, $l_1 = 50$ and $l_2 = 85$. Here, we also compared the proposed tampering detection results with those of the algorithm [26]. Based on the ground truth labels of forged images, Table 9 summarize the performance of these methods in terms of the confusion matrix and F_1 -measure for 10 tampered images. As is obvious in this table, the proposed approach outperforms the algorithm [26] for detecting the tampered regions.

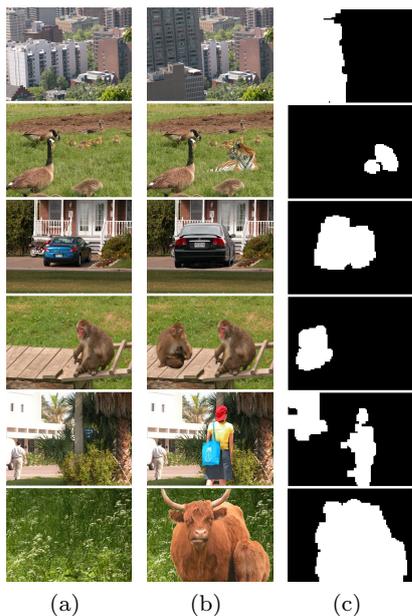


Fig. 12: The application of our method to actual forensic investigations: (a) original JPEG images, (b) tapered images in which the inserted objects have undergone single compression and other regions have borne double compression, and (c) forgery detection results.

Table 7: The performance of the proposed low-dimensional features on CCID database in comparison with other methods in terms of percentage

Method	Recall	Precision	F_1 -measure
Li's method [25]	51.87	99.18	68.12
Milani's method [32]	35.76	96.87	52.23
Dong's method [11]	53.26	98.86	69.23
Proposed approach	56.76	99.08	72.18

Table 8: The performance of various methods in the quantization-semi-aware case on UCID database in %

Method	Recall	Precision	F_1 -measure
Li's method [25]	84.78	99.80	91.68
Milani's method [32]	42.04	99.50	59.11
Dong's method [11]	84.50	99.82	91.52
Proposed approach	86.44	99.67	92.58

Table 9: The performance of the proposed forgery detection approach against the algorithm [26] in percentage

Confusion matrix		Ground truth			
		Algorithm [26]		Our approach	
		C_2	C_1	C_2	C_1
Predicted result	C_2	97.80	45.89	95.58	13.41
	C_1	2.20	54.11	4.42	86.59
F_1 -measure		93.44		96.08	

6 Conclusions and Perspectives

In this paper, we developed a quantization unaware method for double JPEG compression detection. In the

quantization unaware approach, only a global learned model is employed to detect double compression. Experiments demonstrated that not only the proposed algorithm have outstanding performance in the quantization unaware scenario, but also outperformed its competitors in the quantization-semi-aware case. It points out that the proposed model is able to better find out the global structure of single and double compressed images. We also examined the effectiveness of the proposed approach to actual image forgery detection.

Our theoretical study as well as empirical evaluation demonstrate that the approaches like [25, 32, 36, 38] have scarified the telltale footprints lied in the partitions of quantized AC modes. Whereas, the number of quantized AC modes required to double JPEG compression detection is related to the image's block texture and the compression's quality level and all attributes may be informative. Accordingly, a double compression detection algorithm is proposed to leverage the footprints introduced by all AC frequency modes in a low-dimensional representation for achieving an outstanding classification performance. From the proposed statistical investigation, we conclude that, for homogeneous image blocks or low compression's quality levels, the low and middle frequency contents are sufficient to detect double compression. But, for textured patches or the high compression's quality levels, all frequency components are important especially the middle and high frequency bands. As a conclusion, instead of the sparsity in AC modes selection, we suggested the utilization of dimensionality reduction techniques on all single-frequency attributes to, if needed in classification tasks, prevent redundant information and the curse of dimensionality for robust double compression detection.

In another work, we are pursuing the idea of combining the proposed global model with some local information (e.g. a rough estimation of the secondary quality level) of images to improve the performance of our double compression detector.

Acknowledgements The authors would like to thank S. Sabouri for her valuable comments which help us to improve the quality of this paper.

References

1. <http://www.impulseadventure.com/photo/jpeg-quantization.html>.
2. <http://homepages.lboro.ac.uk/~cogs/datasets/ucid/data/ucid.v2.tar.gz>.
3. http://www.shsu.edu/~qx1005/New/Downloads/never_compressed_images.zip.
4. <http://tabby.vision.mcgill.ca>.

5. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
6. Berger, A., Hill, T.P.: A basic theory of Benford's law. *Probability Surveys* **8**, 1–126 (2011)
7. Bianchi, T., Piva, A.: Detection of nonaligned double jpeg compression based on integer periodicity maps. *IEEE Transactions on Information Forensics and Security* **7**(2), 842–848 (2012)
8. Bianchi, T., Piva, A.: Image forgery localization via block-grained analysis of JPEG artifacts. *IEEE Transactions on Information Forensics and Security* **7**(3), 1003–1017 (2012)
9. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)* **2**(3), article 27 (2011)
10. Chen, C., Shi, Y.Q., Su, W.: A machine learning based scheme for double JPEG compression detection. In: *IEEE 19th International Conference on Pattern Recognition (ICPR)*, pp. 1–4 (2008)
11. Dong, L., Kong, X., Wang, B., You, X.: Double compression detection based on Markov model of the first digits of DCT coefficients. In: *IEEE 6th International Conference on Image and Graphics (ICIG)*, pp. 234–237 (2011)
12. Fan, Z., de Queiroz, R.L.: Identification of bitmap compression history: JPEG detection and quantizer estimation. *IEEE Transactions on Image Processing* **12**(2), 230–235 (2003)
13. Farid, H.: Digital image ballistics from JPEG quantization. Tech. Rep. TR2006-583, Department of Computer Science, Dartmouth College (2006)
14. Farid, H.: Exposing digital forgeries from JPEG ghosts. *IEEE Transactions on Information Forensics and Security* **4**(1), 154–160 (2009)
15. Farid, H.: Image forgery detection. *IEEE Signal Processing Magazine* **26**(2), 16–25 (2009)
16. Haralick, R.M., Shanmugam, K., Dinstein, I.H.: Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics SMC-3*(6), 610–621 (1973)
17. Hasimoto-Beltrán, R., Baqai, S., Khokhar, A.: Transform domain inter-block interleaving schemes for robust image and video transmission in ATM networks. *Journal of Visual Communication and Image Representation* **15**(4), 522–547 (2004)
18. He, H., Garcia, E.A.: Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering* **21**(9), 1263–1284 (2009)
19. Huang, F., Huang, J., Shi, Y.Q.: Detecting double JPEG compression with the same quantization matrix. *IEEE Transactions on Information Forensics and Security* **5**(4), 848–856 (2010)
20. Jain, R., Kasturi, R., Schunck, B.G.: *Machine vision*. McGraw-Hill, Inc. (1995)
21. Jolliffe, I.T.: *Principal component analysis*. Springer Series in Statistics, The 2nd edition (2002)
22. Kornblum, J.D.: Using JPEG quantization tables to identify imagery processed by software. The 8th Digital Forensic Research Workshop, *Digital Investigation* **5**, S21–S25 (2008)
23. Lai, S., Böhme, R.: Block convergence in repeated transform coding: JPEG-100 forensics, carbon dating, and tamper detection. In: *IEEE 38th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3028–3032. IEEE (2013)
24. Li, B., Ng, T.T., Li, X., Tan, S., Huang, J.: Revealing the trace of high-quality JPEG compression through quantization noise analysis. *IEEE Transactions on Information Forensics and Security* **10**(3), 558–573 (2015)
25. Li, B., Shi, Y.Q., Huang, J.: Detecting doubly compressed JPEG images by using mode based first digit features. In: *IEEE 10th Workshop on Multimedia Signal Processing*, pp. 730–735 (2008)
26. Li, X.H., Zhao, Y.Q., Liao, M., Shih, F.Y., Shi, Y.Q.: Detection of tampered region for JPEG images by using mode-based first digit features. *EURASIP Journal on advances in signal processing* pp. 1–10 (2012)
27. Lin, Z., He, J., Tang, X., Tang, C.K.: Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis. *Pattern Recognition* **42**(11), 2492–2501 (2009)
28. Liu, Q., Sung, A.H., Qiao, M.: A method to detect JPEG-based double compression. In: *The 8th International Conference on Advances in Neural Networks, Lecture Notes in Computer Science, Part II*, pp. 466–476 (2011)
29. Liu, Q., Sung, A.H., Qiao, M.: Neighboring joint density-based JPEG steganalysis. *ACM Transactions on Intelligent Systems and Technology (TIST)* **2**(2), article 16 (2011)
30. Luo, W., Huang, J., Qiu, G.: JPEG error analysis and its applications to digital image forensics. *IEEE Transactions on Information Forensics and Security* **5**(3), 480–491 (2010)
31. Mahdian, B., Saic, S.: Detecting double compressed JPEG images. In: *IEEE 3rd International Conference on Crime Detection and Prevention*, pp. 1–6 (2009)
32. Milani, S., Tagliasacchi, M., Tubaro, S.: Discriminating multiple JPEG compressions using first digit features. In: *IEEE 37th International Conference on Acoustics, Speech and Signal Processing*

- (ICASSP), pp. 2253–2256 (2012)
33. Narayanan, G., Shi, Y.Q.: A statistical model for quantized AC block DCT coefficients in JPEG compression and its application to detecting potential compression history in bitmap images. In: The 9th International Workshop on Digital Watermarking, Lecture Notes in Computer Science, vol. 6526, pp. 75–89 (2011)
 34. Olmos, A., Kingdom, F.A.A.: A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception* **33**(12), 1463–1473 (2004)
 35. Pérez-González, F., Heileman, G.L., Abdallah, C.T.: Benford’s law in image processing. In: IEEE International Conference on Image Processing (ICIP), vol. 1, pp. I–405–I–408 (2007)
 36. Pevný, T., Fridrich, J.: Detection of double-compression in JPEG images for applications in steganography. *IEEE Transactions on Information Forensics and Security* **3**(2), 247–258 (2008)
 37. Piva, A.: An overview on image forensics. *ISRN Signal Processing* p. article ID 496701 (2013)
 38. Popescu, A.C., Farid, H.: Statistical tools for digital forensics. In: The 6th International Workshop on Information Hiding, Lecture Notes in Computer Science, vol. 3200, pp. 128–147 (2005)
 39. Schaefer, G., Stich, M.: UCID: an uncompressed color image database. In: Storage and Retrieval Methods and Applications for Multimedia, vol. 5307, pp. 472–480 (2004)
 40. Sencar, H.T., Memon, N.: Identification and recovery of JPEG files with missing fragments. *Digital Investigation* **6**, S88–S98 (2009)
 41. Taimori, A., Razzazi, F., Behrad, A., Ahmadi, A., Babaie-Zadeh, M.: A proper transform for satisfying Benford’s law and its application to double JPEG image forensics. In: IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 000,240–000,244 (2012)
 42. Wallace, G.K.: The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics* **38**(1), xviii–xxxiv (1992)
 43. Zach, F., Riess, C., Angelopoulou, E.: Automated image forgery detection through classification of JPEG ghosts. In: Proceedings of Joint 34th DAGM and 36th OAGM Symposium, Lecture Notes in Computer Science, vol. 7476, pp. 185–194 (2012)