Noname manuscript No.
(will be inserted by the editor)

A Part-Level Learning Strategy for JPEG Image Recompression Detection

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

Received: date / Accepted: date

Abstract Recompression is a prevalent form of multimedia content manipulation. Different approaches have been developed to detect this kind of alteration for digital images of well-known JPEG format. However, they are either limited in performance or complex. These problems may arise from different quality level options of JPEG compression standard and their combinations after recompression. Inspired from semantic and perceptual analyses, in this paper, we suggest a part-level middle-out learning strategy to detect double compression via an architecturally efficient classifier. We first demonstrate that singly and doubly compressed data with different JPEG coder settings lie in a feature space representation as a limited number of coherent clusters, called parts. To show this, we visualize behavior of a set of prominent Benford-based features. Then, by leveraging such discovered knowledge, we model the issue of double JPEG compression detection in the family of feature engineering-based approaches as a part-level classification problem to cover all possible JPEG quality level combinations. The proposed strategy exhibits low complexity and

*Ali Taimori (corresponding author) and Farbod Razzazi
Department of Electrical and Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran 14778-93855, Iran
E-mail: alitaimori@yahoo.com, 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

yet comparable performance in comparison to related methods in that family. For reproducibility, our codes are available upon request to fellows.

Keywords Double JPEG compression · information visualization · middle-out process · part-based detection · quality level · sparse coding

1 Introduction

JPEG images are widely used in today's digital photography [10]. Consequently, huge numbers of photo manipulations deal with this prominent image coding scheme. One of prevalent manipulation forms is to load an original camera-captured JPEG image in a photo-editing software tool and after doing some alterations, recompress it in the same JPEG format. This process generates a doubly compressed JPEG image. Therefore, blind detection of double JPEG compression is of great importance among researchers in digital forensics community [7, 22].

So far, three general approaches have been developed for detecting recompression, namely feature engineering- and deep learning-based families, and mixed models of them. Feature engineering-based methods employ forensic experts knowledge to find out discriminative hand-crafted features [2, 6, 13, 18, 20, 23, 25–27]. However, deep learning-based approaches rely on extracting features directly and automatically from image data such as Convolutional Neural Networks (CNNs) [1, 12]. By gaining benefits of the former methods, the third family generally feeds elaborated characteristics to CNNs [14]. Here, we focus on the first family of recompression detectors.

Historically, the problem of double compression detection has been first addressed by Lukáš and Fridrich [18], and Popescu and Farid [23]. Afterwards, Li et al. work [13] was a pioneer in developing modern double compression detection approaches based on the first significant digit characteristics of individual Alternating Current (AC) modes, called Benford-based features here. In [20], Milani et al. suggested a multiple compression detection scheme based on some discriminative significant digits. By leveraging inter-block correlation of quantized DCT coefficients, Dong et al. devised a double compression detector based on Markov modeling [6]. Due to performance and applicability limitations, other techniques were also developed [2, 25, 26]. For instance, in [26], we proposed two Top-Down (TD) and Bottom-Up (BU) learning strategies to discriminate between singly compressed images and doubly compressed ones. In the TD model, singly and doubly compressed data are coarsely categorized into two groups. In contrast, considering intra-class variability, the BU strategy assigns a specific class for each singly or doubly compressed image based on their compression quality factors to be able to both detect recompression and estimate the adjusted quality factors and then their related quantization tables during image acquisition or manipulation processes.

Although the BU approach has the inherent potential to estimate quality levels, it as well as other similar methods such as [6, 13, 20] are not optimal for merely detecting double compression trace, due to their high complexity

or low performance issues for considering all possible combinations of quality level in JPEG coder. Therefore, the field of forensics needs tools for handling quality level options in JPEG standard. In this regard, our contributions are: 1) statistical analysis of behavior of original and recompressed JPEG images with different quality factors, 2) extraction of concept of part from analyzed data as a mild learning strategy for recompression detection, and 3) design of an architecturally efficient part-level classifier for detecting double compression.

The remainder of the paper is organized as: Section 2 describes our proposed Middle-Out (MO) process for identifying parts. Section 3 suggests a part-level learning strategy for recompression detection. Section 4 provides experiments for evaluating both performance and complexity of our architecture and competing ones. The paper is concluded in Section 5.

2 Parts identification

In this paper, a part-level approach, called middle-out learning strategy, is proposed for detecting image recompression. Our basic idea stems from perceptual sub-semantic learning mechanism [5, 8, 24, 28]. In this context, part-based deformable/rigid models are a form of sub-semantic classifiers [5, 8]. In a generic part-based object detector, a set of mid-level describable components are employed for discriminating meaningful object's parts in an image. For instance, to detect a car at a traffic scene, one can individually learn headlights, grill, tire, and license plate components. Then, in detection phase, decision is made based on the presence/absence of trained parts in the query image. Such interesting models are intrinsically resilient to intra-class, view and photometric variations, partial occlusion, and geometric deformations. We borrow the concept of part detection for forensic investigations.

In the proposed part-level middle-out recompression detector, we have analyzed intra-class variations problem of a viewpoint different from TD and BU approaches [25, 26]. By analyzing structures of singly and doubly compressed image data with different quality levels, we have found coherency among some data. The coherency is related to a comparison between the primary and the secondary quality levels used in recompressed JPEG images, where double compression group can be semantically partitioned into three sub-sets, namely cases in which the first compression's quality factor is less than, equal to, or more than the second one. At a glance, the MO strategy for detecting double compression provides three main benefits: 1) middle-out detector is simple and has yet similar performance in comparison to the sophisticated BU process; 2) the proposed model does not need defining many classes for identifying recompression of different JPEG coder settings, which brings generalization capability in practice; and 3) the middle-out learning strategy can intrinsically predict the history of recompression from aspects of applied successive compression's quality factors during coding. To find the coherency, in below, we first demonstrate by a hypothesis that structures of singly and doubly compressed data in a feature space representation can be partitioned into four

distinct clusters via an information visualization technique. Then, by exploiting the information, we model the strategy of learning for double compression detection as a 4-class classification problem in Section 3.

Definition 1 (Single/double compression parts) *Assume $l_1 \in \mathcal{Q}$ and $l_2 \in \mathcal{Q}$ are respectively the first and the second compression's quality levels of singly and doubly compressed JPEG images, in which $\mathcal{Q} = \{1,\ 2,\ \cdots,\ 100\}$. Then, those singly compressed images with the quality factor l_1 as well as doubly compressed images having the compression history of $l_1 < l_2$, $l_1 = l_2$, and $l_1 > l_2$ are respectively defined as parts P_1, P_2, P_3, and P_4. Figure 1 (a) schematizes the three-level hierarchy of the proposed middle-out learning strategy in order to detect double compression for the frequently used set of $\mathcal{Q}^S \subset \mathcal{Q}$ by researchers, but not limited to that, where $\mathcal{Q}^S = \{50,\ 55,\ \cdots,\ 95\}$.*

Hypothesis 1 (Partitionability of single/double compression parts) *Let the matrix $\mathbf{L} \triangleq \{\mathbf{v}_i\}_{i=1}^{N}$ contain initial features belonging to single and double compression parts. Then, we empirically demonstrate that four single and double compression parts defined in Definition 1 have distinct coherent clusters in the 2-D scatterplot and are partitionable.*

Test: To statistically examine validity of Hypothesis 1 on a data set, we utilized Benford-based features proposed in [25] and feature visualizer of [26]. As shown in Fig. 1 (b), feature extractor of [25] partitions each non-overlapping 8×8 quantized DCT blocks of JPEG data into one Direct Current (DC) mode in the top-left position of grid ($\mathcal{D}$) and 63 zig-zag AC modes ($\mathcal{A}_1$ to $\mathcal{A}_{630}$). Because content of the DC mode is not information rich, it is discarded. Therefore, for each AC mode of a questionable image, probability mass functions pertaining to the first significant digits from 1 to 9 as well the digit 0 from the second significant digit are determined as Benford-based features. This process leads to the feature dimension of $m = (64-1) \times (9+1) = 630$ for each image. The visualizer of these features consists of a bi-level dimensionality reduction.

At the first level of visualizer, the original features are sparsely coded by compressed sensing technique and group Least Absolute Shrinkage and Selection Operator (LASSO) algorithm [4, 15]. To this intent, we randomly choose learning and testing sets with replacement from a given singly and doubly compressed images data set. We construct a $m \times n$ over-complete dictionary as $\mathbf{\Psi} = \left\{ \mathbf{\Psi}_1\ \mathbf{\Psi}_2\ \mathbf{\Psi}_3\ \mathbf{\Psi}_4 \right\}$ for which $\mathbf{\Psi}_j \in \mathbb{R}^{m \times T_l},\ \forall j = 1, \cdots, 4$, and $n = 4T_l$. The parameter T_l denotes the number of learning images for each part. For the sample $\mathbf{x}_i \in \mathbb{R}^m$ drawn from the testing set, atomic decomposition is done by $\mathbf{x}_i = \mathbf{\Psi}\boldsymbol{\alpha}_i,\ \forall i$. We determine the estimated sparse vector $\widehat{\boldsymbol{\alpha}}_i$ by solving the group LASSO as

$$\widehat{\boldsymbol{\alpha}}_i = \underset{\boldsymbol{\alpha}_i}{\operatorname{argmin}} \left\{ \frac{1}{2} \left\| \mathbf{\Phi}\mathbf{\Psi}\boldsymbol{\alpha}_i - \mathbf{\Phi}\mathbf{x}_i \right\|_{L_2}^{2} + \lambda \left\| \boldsymbol{\alpha}_i \right\|_{L_2/L_1} \right\}, \tag{1}$$

in which the measurement matrix $\mathbf{\Phi} \in \mathbb{R}^{d \times m}$ is a Gaussian random matrix with the distribution $\mathcal{N}(0,\ 1)$ and the mixed L_2/L_1-norm is defined as $\left\| \boldsymbol{\alpha}_i \right\|_{L_2/L_1} \triangleq$

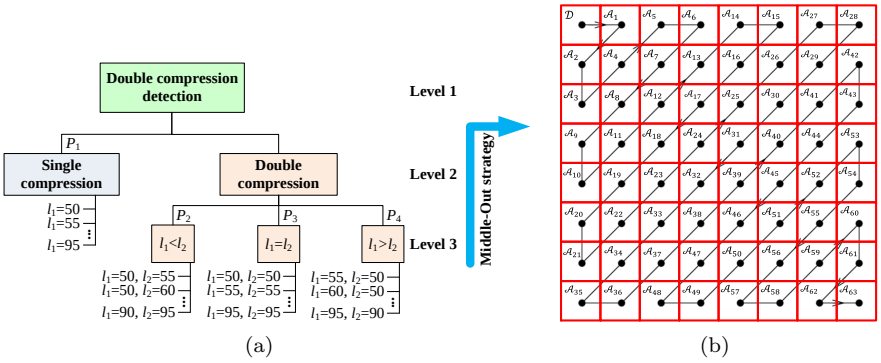


Fig. 1: (a) The hierarchy of middle-out parts learning of P_1 to P_4 for detecting double compression. Parameters l_1 and l_2 denote the first and the second compression's quality levels, respectively, and (b) single DC mode of $\mathcal{D}$ and AC modes of $\mathcal{A}_1$ to $\mathcal{A}_{63}$ in zig-zag order defined in JPEG compression standard.

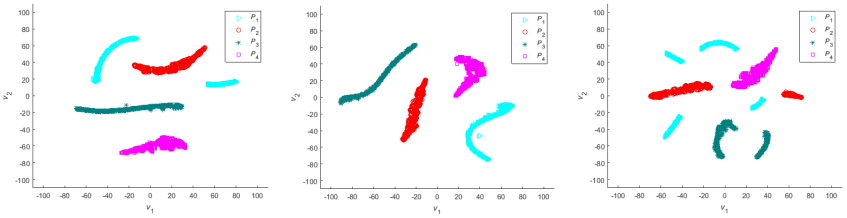


Fig. 2: From (a) to (c), test results of Hypothesis 1 under 3 random runs of the utilized visualizer. Spline-like clusters satisfy partitionability of single/double compression parts of P_1 to P_4.

$\sum_{k=1}^{4} \left\| \boldsymbol{\alpha}_i^k \right\|_{L_2}$, for which $\boldsymbol{\alpha}_i^k \triangleq \left\{ \boldsymbol{\alpha}_j^i \mid j \in k \right\}$. The coefficient $\lambda \in [0,\ 1]$ denotes a regularization parameter tuned experimentally as $\lambda = 0.5$ and the parameter d is the number of measurements which satisfies the condition $d < m = 630$. We selected $d = 200$, which considerably reduces computational complexity. Sparse vectors are placed in the matrix $\mathbf{D} \in \mathbb{R}^{n \times N}$, where $N = 4T_t$. The variables n and T_t represent the number of atoms in the dictionary and testing images for each class, respectively. We set $T_l = T_t$ and changed them from 200 to 3000, so that its sub-optimal value was 450 to appropriately visualize partitionability of parts.

At the second level of the visualizer, the matrix containing estimated sparse vectors are mapped onto the 2-D visual representation matrix of $\mathbf{L} \triangleq \{\mathbf{v}_i\}_{i=1}^{N}$, where $\mathbf{v}_i = \left[v_1^i \ v_2^i \right]^{\mathrm{T}}$, $\forall i$, using the unsupervised non-parametric t-SNE dimensionality reduction algorithm of [19]. Figure 2 illustrates test results of Hypothesis 1 under 3 random runs. The plots reveal that each part is compactly clustered in the 2-D feature space and to some extent satisfy partition-

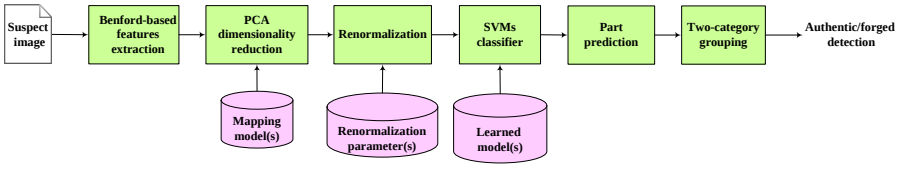


Fig. 3: The suggested part-level learning approach in a pipeline system.

ability of single/double compression parts. The point clouds of parts P_1 to P_4 are pertaining to $4 \times 450 = 1800$ singly and doubly compressed JPEG images created by original raw McGill CCID image-set [21]. It is noticeable that our examinations on other feature sets such as those proposed in [6, 13, 20] demonstrate that Hypothesis 1 is yet valid. ■

3 Part-level learning for recompression detection

Statistical analysis done in Section 2 shows singly and doubly compressed data with different JPEG quality factor settings lie in the feature space as a limited number of coherent clusters, called parts. We leverage this discovered knowledge to model the issue of double JPEG compression detection in a family of feature engineering-based approaches as a part-level classification problem to cover all possible combinations of the JPEG quality level. To do this, we consider each part as a distinct mid-level object for classification. This leads a mild learning architecture, where we have only 4 classes for all cases of the quality levels l_1 and l_2. We used SVMs classifier with linear kernel for learning a 4-class classification problem. Because SVMs are basically designed for binary classification, we utilized one vs. one division strategy to handle the multi-class scenario [11]. To identify recompressed data, we finally group the predicted part P_1 as single compression class and all the classified parts P_2, P_3, and P_4 as double compression class.

Figure 3 depicts a pipeline system for recompression detection, in which we are able to employ the proposed part-level learning architecture. The detector first receives a suspect image and then extracts its forensic Benford-based features. To handle the curse of dimensionality, we diminish the feature dimension $m = 630$ using a pre-learned PCA model. It is important to note the dimensionality reduction approach embedded in the visualizer cannot be exploited in the proposed learning architecture. In fact, the t-SNE technique is a non-parametric method and thus does not naturally deliver an explicit mapping model to be utilized for pattern classification problems. After applying the PCA, reduced features are renormalized by the vectors of sample mean, $\boldsymbol{\mu}$, and standard deviation, $\boldsymbol{\sigma}$. For the n^{th} learning/testing low-dimensional sample of $\mathbf{y}_n$, normalization in the range of $[0,\ 1]$ is done as

$$\widetilde{\mathbf{y}}_n = (\mathbf{y}_n - \boldsymbol{\mu}) \oslash \boldsymbol{\sigma}, \tag{2}$$

where the symbol $\oslash$ denotes the entry-by-entry division and $\mathbf{y}_n = \mathbf{T}^{\mathrm{T}}\mathbf{x}_n$. The vector $\mathbf{x}_n$ represents Benford-based features of the n^{th} sample. The matrix $\mathbf{T} \in \mathbb{R}^{m \times d_r}$ is PCA transformation kernel related to a given model, in which d_r represents the dimensionality of reduced features. Then, the renormalized features are fed to SVMs classifier for predicting parts. At the end, the part decision is assigned to groups of single and double compression in order to classify authentic singly compressed data from forged doubly compressed contents. It is notable that databases of mapping model(s), renormalization parameter(s), and learned model(s) are obtained from training phase of the proposed approach by all allocated learning ensembles. These predetermined models may be one or more based on forensic scenarios considered in the problem of double compression detection, which will be discussed in detail in the next section (See also Table 3 for more details.).

4 Experiments

4.1 Experimental settings

For training and testing purposes, we employed McGill CCID raw images repository including 1096 TIFF images of dimensions 768×576 or vice versa [21, 26]. In experimental simulation, we created different singly and doubly compressed JPEG images from sampled quality factors in the frequently used set of $\mathcal{Q}^S = \{50, 55, \cdots, 95\}$ by researchers. To do this, at first, we compressed all images of the raw data set with the quality factor $l_1 \in \mathcal{Q}^S$. Then, each singly compressed image was recompressed with the quality factor $l_2 \in \mathcal{Q}^S$. The collection of singly and doubly compressed images forms our utilized data set. We divided the data set into two distinct sets based on hold-out cross validation by separating 75% of them for learning and the remainder 25% for testing. Experiments independently performed on two important forensic scenarios, in which quantization information of a JPEG file may be known or unknown. These scenarios have respectively called quantization-unaware and quantization-semiaware double compression detection [25, 26]. For both scenarios, we empirically set the reduced features dimension to $d_r = 300$. We have also compared our approach with competing ones [6, 13, 20, 26]. Codes were implemented in MATLAB and run on an Intel Core i7 2.2GHz machine. We employed LIBSVM toolbox for implementing SVMs classifier [3].

4.2 Detection under quantization-unaware scenario

According to the discussed methodology of data set creation, the aggregated singly and doubly compressed images will result in an unbalanced part distribution as graphically illustrated in Fig. 4 (a). The imbalance ratio may bias learning of one part against another in the one vs. one architecture employed in SVMs classifier. To alleviate this problem, we added learning data of parts

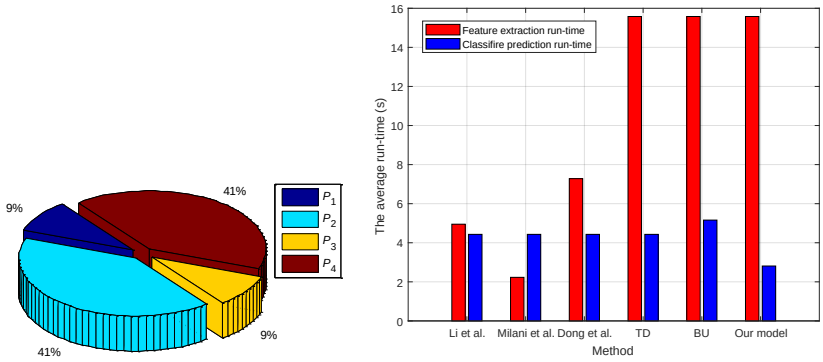
P_1 and P_3 from an auxiliary data set to the corresponding low populated classes. We used NCID data set, which contains 5150 BMP images of dimensions 256×256 [16, 17]. Finally, the number of balanced SVMs learning data for parts P_1 to P_4 was 46850, 36990, 46850, and 36990, respectively.

In evaluation phase, confusion matrix for 4-class prediction of parts was as

$$\mathbf{C} = \begin{pmatrix} 61.72 & 2.13 & 62.01 & 9.39 \\ 0.58 & 94.55 & 0.58 & 1.22 \\ 8.39 & 0.74 & 8.03 & 2.77 \\ 29.31 & 2.58 & 29.38 & 86.63 \end{pmatrix}, \tag{3}$$

where each entry of the matric $\mathbf{C}$ is $c_{ij} \triangleq P(i|j)$ in terms of percentage, $\forall i, j = 1, \cdots, 4$. The conditional probabilities show promising results, except for the part P_3. This is due to the nature of employed Benford-based features which do not cover double compression detection in the same quality factors as well as failure of PCA dimensionality reduction to discriminate between doubly compressed data with $l_1 = l_2$ and the part P_1 as singly compressed ones (See the entry $c_{13} = 62.01\%$ in (3) as misclassification of class P_3 to class P_1.). Factually, in terms of the employed features, this method does not support double compression detection in the case $l_1 = l_2$. We refer readers to [12, 27] for dealing with such a specific scenario. Additionally, the entry $c_{41} = 29.31\%$ narrates more than $\frac{1}{4}$ of authentic camera-captured data is misclassified as double compression class with $l_1 > l_2$. The philosophy behind this phenomenon is basically that singly compressed data with high image quality as well doubly compressed images having $l_1 > l_2$ do not suffer of a considerable information loss at the first level of compression. This may also interpret more discriminative features are required to add to the Benford-based characteristics for grasping higher performance. Note that the focus of current work is on the architecture of classifier rather than feature extraction.

Table 1 reports in detail effect of various quality factors on performance of our detector. The parameters A_{G_1} and A_{G_2} denote detection accuracy percentage for single and double compression groups, respectively. As interpreted above, the accuracy on single compression prediction shows an incremental envelope by growing the quality level; whereas, the situation is just reverse for detecting double compression in the case $l_1 = l_2$. Generally speaking, cost of misclassifying double compression class as single one is more than the inverse case of c_{41} [25]. Therefore, our emphasis hereafter is more on analysis of double compression prediction. In order to save space, we summarize performance of different competing methods in Table 2 based on criteria of Recall, Precision, and F_1-measure. The later is a mixture of Recall and Precision. Obviously, our strategy has the second rank of F_1-measure near to the best complex BU, while the proposed MO method has less complexity, as it will be seen in Section 4.4.



(a) (b)

Fig. 4: (a) The intense unbalanced population from the parts of training set under Quantization-unaware scenario before balancing process, and (b) the average feature extraction and classifier prediction run-times of various learning strategies under quantization-semiaware scenario.

Table 1: Our method's accuracy for various combinations of quality level (%)

A_{G_2}		l_1									
		50	55	60	65	70	75	80	85	90	95
l_2	50	59.85	95.26	83.94	100	100	99.64	93.8	91.24	91.24	87.59
	55	97.45	68.61	94.16	91.97	100	100	89.05	87.96	95.99	45.26
	60	93.07	91.61	44.53	87.96	99.64	100	98.91	71.17	70.8	90.15
	65	100	100	89.42	48.54	98.18	100	100	88.32	97.45	60.95
	70	97.45	96.35	100	90.51	30.29	98.91	100	98.54	78.47	85.04
	75	99.27	99.27	98.18	100	73.72	22.63	87.23	100	90.51	78.83
	80	98.91	98.18	98.54	97.45	100	93.43	35.77	98.54	93.43	87.59
	85	100	99.27	100	99.64	98.54	97.08	100	28.83	100	87.59
	90	100	100	100	100	99.27	99.64	100	99.64	29.93	82.12
	95	99.64	100	100	99.64	100	99.27	100	100	99.64	10.95
A_{G_1}		40.15	31.02	55.47	50.73	70.07	77.74	63.87	71.53	68.61	87.96

4.3 Detection under quantization-semiaware scenario

In quantization-semiaware scenario, an individual model is learned for each quality factor in $\mathcal{Q}^S$. Hence, the total number of learned models in this case will be equal to 10 (See Table 3 for more details.). Based on the last quality level available in JPEG file, population of each part for learning doubly compressed data is variable. Also, for boundary quality factors, namely 50 or 95 in $\mathcal{Q}^S$, the number of double compression parts is reduced to 2. This means the detection in these cases is simplified to a 3-class classification problem. The performance metrics for compared methods are summarized in Table 2. As it is shown, the proposed strategy exposes a comparable performance than the sophisticated BU approach. As a generalization test, we also fed other input

Table 2: Performance comparison in two different scenarios (%)

Method	Quantization-unaware scenario						Quantization-semiaware scenario					
	Recall				Precision	F_1-measure	Recall				Precision	F_1-measure
	P_2	P_3	P_4	Total			P_2	P_3	P_4	Total		
Ref. [13]	84.96	3.58	28.05	51.21	99.31	67.58	99.77	2.96	96.63	88.68	99.68	93.86
Ref. [20]	51.56	8.98	19.79	33.00	97.32	49.29	95.17	11.72	65.19	73.34	98.42	84.05
Ref. [6]	85.39	4.38	29.40	52.09	99.17	68.31	99.84	3.69	96.33	88.65	99.59	93.8
TD [26]	84.91	5.73	44.03	58.59	99.03	73.62	99.76	2.41	97.65	89.08	99.73	94.1
BU [26]	99.8	47.15	99.02	94.18	95.08	94.63	99.93	47.15	99.64	94.52	95.19	94.85
Proposed method	97.87	37.99	90.61	88.61	95.86	92.09	99.81	52.01	98.85	94.59	94.61	94.6

features such as those proposed in [6, 13, 20] to our classifier. For the feature sets of [6, 13], similar classification performance were obtained. Under this setting, MO+features of [20] failed to detect singly compressed data correctly. By comparing the metric F_1-measure between different scenarios of Table 2, we find out that the problem in the quantization-semiaware case is simpler than the quantization-unaware scenario.

4.4 Complexity analysis

Here, we analyze complexity of [13], [20], [6], TD and BU [26], and the suggested MO approaches for detecting double compression from a JPEG image with compression's quality of $l \in \mathcal{Q}$. Various parameters may be effective for analyzing complexity of a recompression detector such as the number of possible combinations for l_1 and l_2 as a measure of the number of species, the feature dimension, the number of classes, classifier's complexity (e.g., the number of support vectors in SVMs), and memory requirements (e.g., the size of learned models on storage space). Our intent is to theoretically analyze some of these criteria in an independent manner. For instance, in quantization-unaware scenario, the BU process requires $\mathcal{O}\left(|\mathcal{Q}| + |\mathcal{Q}| \cdot |\mathcal{Q}|\right) \approx \mathcal{O}\left(|\mathcal{Q}|^2\right)$ class to detect recompression regardless of other underlying factors. The symbol $|\cdot|$ denotes the cardinality of a set. Such an order is approximately equal to $\mathcal{O}\left(|\mathcal{Q}|\right)$ for quantization-semiaware scenario; whereas our middle-out strategy, similar to the TD approach, performs at a constant order independent from the size of $\mathcal{Q}$ in both scenarios. Table 3 tabulates a comparison of the complexity of different learning strategies for the utilized set $\mathcal{Q} = \mathcal{Q}^S$. A case by case comparison shows that our proposed learning architecture has a mid complexity than [13], [20], [6], TD, and BU methods.

Practically, by considering all computational complexity factors, Fig. 4 (b) shows real average run-time of competing approaches consisting of feature extraction and classifier prediction run-times under quantization-semiaware scenario. Feature dimension for [13], [6], [20], TD and BU methods [26], and our model is 180, 27, 200, 300, 300, 300, respectively. This means three later approaches take more time for feature extraction, as justified in Fig. 4 (b). In terms of classifier prediction run-time and performance trade-off, the proposed

Table 3: The complexity of various learning strategies

Method	Forensic scenario	Complexity in terms of $\mathcal{Q}$	Number of classes for $\mathcal{Q}-\mathcal{Q}^S$	Number of learned models for $\mathcal{Q}-\mathcal{Q}^S$
Refs. [6, 13, 20], TD [26]	Quantization-unaware	$\mathcal{O}(1)$	2	1
	Quantization-semiaware	$\mathcal{O}(1)$	2	$\lvert\mathcal{Q}^S\rvert - 10$
BU [26]	Quantization-unaware	$\mathcal{O}(\lvert\mathcal{Q}\rvert^2) \approx 10000$	$\lvert\mathcal{Q}^S\rvert(1+\lvert\mathcal{Q}^S\rvert) - 110$	1
	Quantization-semiaware	$\mathcal{O}(\lvert\mathcal{Q}\rvert) - 100$	$1+\lvert\mathcal{Q}^S\rvert - 11$	$\lvert\mathcal{Q}^S\rvert - 10$
Proposed method	Quantization-unaware	$\mathcal{O}(1)$	4	1
	Quantization-semiaware	$\mathcal{O}(1)$	3 or 4[†]	$\lvert\mathcal{Q}^S\rvert - 10$

[†] For boundary quality factors, the number of classes is 3. Otherwise, it is equal to 4.

learning architecture exhibits a run-time about half of the high performance BU approach without sacrificing the F_1-measure performance considerably, as reported in Table 2. The reason behind the reduction in F_1-measure of our classifier in comparison to [26] is that, the number of classes in [26] grows with the number of combinations of the quality levels of l_1 and l_2 as shown in Table 3, which interprets considering more details for classification. However, the proposed middle-out classifier uses only 3 or 4 classes. This fact brings appropriate simplicity in cost of an inevitable decay in F_1-measure.

5 Conclusion

This paper introduces a simple and yet efficient JPEG recompression detector. In the proposed part-level architecture, we statistically demonstrated in a complicated, unaware scenario of quantization information that defined parts have semantic concepts in data structure of single and double compression. The concept of parts originate from situation of compression's quality factors adjusted during JPEG coding. In addition to single and double compression detection capability, our middle-out learning strategy can inherently predict the history of recompression from aspects of applied successive compression's quality factors during coding, i.e. parts pertaining to settings of $l_1 < l_2$, $l_1 = l_2$, and $l_1 > l_2$ (The parameters l_1 and l_2 represent the first and the second compression's quality levels of singly and doubly compressed JPEG images, respectively.). This ability of the suggested method may be useful in forensic identification of source devices [7, 9, 10, 22]. Also, incorporating the limited parts in our model brings a generalization capability of dealing with various commercial JPEG coders. As an intermediate-level learning architecture, the proposed model can perceptually play a complementary role for top-down and bottom-up learners of [26]. Hence, as a future study, we are looking for a perceptual fused version of complementary top-down, bottom-up, and middle-out learning strategies to synergistically grasp court-friendly performance. We also intend to add a set of features based on sparse coding for more discrimination and to be able to detect double compression in the case $l_1 = l_2$.

References

1. Alipour N, Behrad A (2020) Semantic segmentation of JPEG blocks using a deep CNN for non-aligned JPEG forgery detection and localization. Multimed Tools Appl pp 1–17

2. Bianchi T, Piva A (2012) Image forgery localization via block-grained analysis of JPEG artifacts. IEEE Trans Inf Forensics Security 7(3):1003–1017
3. Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. ACM T Intel Syst Tec 2(3), article 27
4. Cheng H, Liu Z, Yang L, Chen X (2013) Sparse representation and learning in visual recognition: theory and applications. Signal Process 93(6):1408–1425
5. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: IEEE CVPR, vol 1, pp 886–893
6. Dong L, Kong X, Wang B, You X (2011) Double compression detection based on Markov model of the first digits of DCT coefficients. In: IEEE ICIG, pp 234–237
7. Farid H (2009) Image forgery detection. IEEE Signal Process Mag 26(2):16–25
8. Felzenszwalb PF, Girshick RB, McAllester D, Ramanan D (2010) Object detection with discriminatively trained part-based models. IEEE Trans Pattern Anal Mach Intell 32(9):1627–1645
9. Fridrich J (2009) Digital image forensics. IEEE Signal Process Mag 26(2):26–37
10. Galvan F, Puglisi G, Bruna A, Battiato S (2014) First quantization matrix estimation from double compressed JPEG images. IEEE Trans Inf Forensics Security 9(8):1299–1310
11. Hsu CW, Lin CJ (2002) A comparison of methods for multiclass support vector machines. IEEE Trans Neural Netw 13(2):415–425
12. Huang X, Wang S, Liu G (2018) Detecting double JPEG compression with same quantization matrix based on dense CNN feature. In: IEEE ICIP, pp 3813–3817
13. Li B, Shi YQ, Huang J (2008) Detecting doubly compressed JPEG images by using mode based first digit features. In: IEEE MMSP, pp 730–735
14. Li B, Zhang H, Luo H, Tan S (2019) Detecting double JPEG compression and its related anti-forensic operations with CNN. Multimed Tools Appl 78(7):8577–8601
15. Liu J, Ji S, Ye J (2011) SLEP: sparse learning with efficient projections. Arizona State University
16. Liu Q, Sung AH, Qiao M (2011) A method to detect JPEG-based double compression. In: ICANN, Lect. Notes Comput. Sc., Part II, pp 466–476
17. Liu Q, Sung AH, Qiao M (2011) Neighboring joint density-based JPEG steganalysis. ACM T Intel Syst Tec 2(2), article 16
18. Lukáš J, Fridrich J (2003) Estimation of primary quantization matrix in double compressed JPEG images. In: DFRWS
19. van der Maaten L, Hinton G (2008) Visualizing data using t-SNE. J Mach Learn Res 9:2579–2605
20. Milani S, Tagliasacchi M, Tubaro S (2014) Discriminating multiple JPEG compressions using first digit features. APSIPA Trans Signal Inf Process 3
21. Olmos A, Kingdom FAA (2004) A biologically inspired algorithm for the recovery of shading and reflectance images. Perception 33(12):1463–1473
22. Piva A (2013) An overview on image forensics. ISRN Signal Process Article ID 496701
23. Popescu AC, Farid H (2004) Statistical tools for digital forensics. In: IWIH, pp 128–147
24. Richtsfeld A, Mörwald T, Prankl J, Zillich M, Vincze M (2014) Learning of perceptual grouping for object segmentation on RGB-D data. J Vis Commun Image R 25(1):64–73
25. Taimori A, Razzazi F, Behrad A, Ahmadi A, Babaie-Zadeh M (2016) Quantization-unaware double JPEG compression detection. J Math Imaging Vis 54(3):269–286
26. Taimori A, Razzazi F, Behrad A, Ahmadi A, Babaie-Zadeh M (2017) A novel forensic image analysis tool for discovering double JPEG compression clues. Multimed Tools Appl 76(6):7749–7783
27. Yang J, Xie J, Zhu G, Kwong S, Shi YQ (2014) An effective method for detecting double JPEG compression with the same quantization matrix. IEEE Trans Inf Forensics Security 9(11):1933–1942
28. Zhang C, Cheng J, Liu J, Pang J, Liang C, Huang Q, Tian Q (2014) Object categorization in sub-semantic space. Neurocomputing 142:248–255