Splicing Forgeries Localization through the Use of First Digit Features

Irene Amerini ♣, Rudy Becarelli ♣, Roberto Caldelli ♣♢, Andrea Del Mastio ♣

♣ *Media Integration and Communication Center - MICC, University of Florence, Florence, Italy*
♢ *National Interuniversity Consortium for Telecommunications - CNIT, Florence, Italy*
{irene.amerini,rudy.becarelli,roberto.caldelli,andrea.delmastio}@unifi.it

***Abstract*—One of the principal problems in image forensics is determining if a particular image is authentic or not and, if manipulated, to localize which parts have been altered. In fact, localization is basic within the process of image examination because it permits to link the modified zone with the corresponding image area and, above all, with the meaning of it. Forensic instruments dealing with copy-move manipulation quite always provides a localization map, but, on the contrary, only a few tools, devised to detect a splicing operation, are able to give information about localization too. In this paper, a method to distinguish and then localize a single and a double JPEG compression in portions of an image through the use of the DCT coefficients first digit features and employing a Support Vector Machine (SVM) classifier is proposed. Experimental results and a comparison with a state-of-the-art technique are provided to witness the performances offered by the proposed method in terms of forgery localization.**

***Index Terms*—Forgery detection, splicing attack, localization, SVM.**

I. Introduction

Looking at an image often raises a question, is it original or has it been retouched? Such a question is usually due to the well-known easiness with which digital images can be modified to alter their content and the meaning of what is represented in them. The context in which pictures are involved could be a tabloid, an advertising poster or an insurance practice, but, also a court of law where images are presented as basic evidences for a trial to influence the judgement; so answering reliably to such questions about integrity becomes fundamental. Image forensics deals with these issues by developing technological instruments which generally allow to determine, only on the basis of a photograph, if that asset has been modified. Furthermore, it would be interesting, once established that something has happened, to understand what: if an object or a person has been covered, if something has been copied from another image localizing the tampering. Regarding forgeries individuation three are the principal classes of detectors studied so far: those based on double JPEG compression [1]–[4] adopted to reveal splicing attack, those based on inconsistent shadows [5] and finally those based on local features descriptors (mainly SIFT - Scale Invariant Feature Transform) [6], [7] used to get rid of copy-move attack. A complete overview of forensic methods devoted to tampering detection is well underlined in [8].
In detail, in this paper the splicing attack has been considered i.e. the case in which a part of an image is grabbed, possibly then adapted (geometrically transformed and/or enhanced) and finally pasted onto another one to build a new fake image. The objective of this paper is to overcome the limitation of the work done in [4], where multi-JPEG compressions are evidenced only in a full-frame image. The proposed method is able to distinguish and then localize a single and a double JPEG compression also in small portions of an image through the use of the DCT coeffcients first digit features and by employing a training/testing procedure using the Support Vector Machine (SVM) classifier. The proposed technique will be also compared with a different method [2], based on the analysis on the JPEG artifacts introduced by the double JPEG compression where an automatic algorithm to localize the forgery is designed.

The rest of this paper is organized as follows. In Section 2, an overview on the related works is done and then in Section 3 we discuss the proposed framework. The Section 4 contains experimental results, while conclusions are finally drawn in Section 5.

II. Related Works

In this Section previous methods, devoted to discriminate forgery and original images, are underlined. Identifying the double JPEG compression in an image may be an initial step to detect image forgeries. In this paper we focus on JPEG images since this kind of compression is the most widely adopted compression standard. Previous methods on discriminating double compressed JPEG images from single compressed JPEG images studied how the shape of the histogram of the DCT coefficient is changed by the secondary JPEG compression. Popescu et al. [9] presented a detection method based on a periodicity measure, which evaluate whether the Fourier transform of the histogram of the quantized DCT coefficients has certain artifacts. In the papers by Fu et al. and Li et al. ([3] and [10] respectively), it is evidenced that double compression would cause the distribution of the first digits of the DCT coefficients violating the generalized Benford's law distribution. In particular, Li et al. proposed to use the probabilities of the first digits of QDCT coefficients from individual modes as features to detect double compressed images. Then with a two-class classification strategy double and single compressed JPEG images can be differentiated.

978-1-4799-8882-2/14/$31.00 ©2014 IEEE

The method in [4], on the other side, relies on a set of SVMs classifiers and allows to identify the number of compression steps applied to an image studying how the distribution of the first significant digits of DCT coefficients is changed. Many works on splicing detection are based on introducing a set of features, deriving from the fact that re-compression induces periodic artifacts and discontinuities in the image histogram, to train a SVM classifier often omitting the information about the forgery localization [11], [12]. Furthermore, in the paper by Bianchi et al. [2] a method to discriminate between original and tampered regions in JPEG images is proposed checking out a double compression (align or not-align) in the image. The algorithm computes a likelihood map indicating the probability for each 8×8 DCT block of being double compressed. To the best of our knowledge very few forensic algorithms (i.e. [1] and [2]) have been designed to localize in an automatic way the tampered regions, so our paper is addressed to tackle such an issue. In fact, in this work, we designed a method able to localize tampering (single and multiple forgeries with various dimensions) in JPEG images, without any prior information about the location of the manipulated area by using a technique based on SVM and exploiting the first digit features proposed in [4]. It is the first time that a method based on Benford's law is employed to solve the problem of tampering localization since only the full-frame image has been taken in account so far. This method is also independent from the dimension of the forgery due to the fact that each not-overlapped block is considered at a time on the contrary to what happens in [2] where the statistics on each block is calculated taking into consideration the full-frame image properties. Furthermore a wide range of first/second step JPEG compressions (quality factor from 50 to 95) has been taken in account in the experiments to demonstrate the validity of our method.

III. Proposed method

In this Section, the proposed method is described: firstly, the theoretical background is briefly recalled in subsection III-A and, then, in subsection III-B, the whole procedure to detect and localize image forgery is introduced.

A. First Digit Features

The Benford's Law (also known as the First Digit Law) is a well known rule in statistics of natural phenomena. According to it, the frequency of appearance of each digit in the first significant place of quantities observed from natural phenomena is logarithmic. As stated in Equation (1), in natural phenomena it arises that:

$$P(d) = log_{10}(1 + \frac{1}{d}) \quad (1)$$

stated that d is the $first\ significant\ digit$ of the measure of the natural phenomenon. What is interesting for us is that in [13] is discussed that even natural images, under certain constraints, follow closely this law. Based on this theory and following the approach stated in [4], we have considered the first digit tendency of the de-quantized DCT (8×8 block) coefficients of an image. When an image is compressed, such a trend is perturbed and the introduced perturbation is different according to the number of occurred compressions. This permits to estimate the number of subsequent compressions an image has suffered, by analyzing the divergence of the DCT first digit statistic with respect to the theoretical Benford's trend.

In this work, we addressed our investigations to single and double compression with the intent to discern between image areas that have and have not been involved in a splicing attack. To do so, given an image, DCT coefficients are extracted and, for each 8×8 block, the first 9 spatial frequencies in zig-zag scan order, with the exception of the DC one [14], are selected. For each spatial frequency, the histogram representing the occurrences of the first digits ($m = 1, 2, \ldots, 9$) is constructed but only the values corresponding to $m = 2, 5, 7$ are taken as distinctive features in order to reduce the dimensionality of the feature vector without loosing significantly in detection performances (see in [4] for such a choice). Finally, the feature vector has a size of 27 (9 DCT coefficients $\times 3$ first digits). The feature vector is then used to train a SVM classifier, in order to distinguish between a positive and a negative set which represents the single and double compressed images respectively. In particular, an array of SVM classifiers is trained, each of them corresponding at the different values of the second compression quality factor (a step of 5 has been considered in the experimental tests, see subsection IV-A).

B. Proposed forgery localization algorithm

To achieve the aim of localization, the SVM has been trained by means of image portions whose size was compliant with that of search window used for forgery localization. In fact, in the first step of the proposed method, the to-be-checked image I ($M \times N$) is divided into non-overlapping blocks B_k whose size is $W \times W$; such a dimension represents the forgery localization resolution and, basically, it is a trade-off between the precision with which a manipulation can be determined eventually and the need that the first digit statistics still holds. This second point becomes crucial especially when a block contains a flat area of the image which consequently causes that almost all AC-DCT coefficients are zero and obviously their first digits have no occurrences at all. Such a critical situation can also happen after a heavy JPEG compression (low *QF* values) which implies that most of AC-DCT coefficients have been nullified. In these circumstances, though the subsequent SVM classifier will provide an answer, it would be wiser not to give a classification of that block: some hints on this issue of decidability will be given within section V.
According to what described in subsection III-A, a 27-dimensional feature vector is computed for each block B_k and passed to the SVM classifier that provides, for each block B_k, a confidence value D_k (it is a signed value) that takes into account of the distance from the secant hyperplane. Usually, the SVM classifier compares such a confidence value with a threshold $T = 0$ and determines if a certain block has been JPEG compressed once or twice. We have decided to exploit such a value to obtain a measure for each block of

the image in order to construct a sort of reliability map at block resolution. Doing so, each block is evaluated as a sub-image independently from the full-frame context and hence its feature descriptor is not affected by the dimension of the whole forged patch.

IV. Experimental Results

In this Section some experimental results are presented in order to validate the proposed work. First, the image datasets, along with their usage and characteristics, are presented together with a description of how SVM training phase has been carried out. In the following subsections, results which provide a qualitative overview of the proposed method performances are introduced, then a quantitative analysis in terms of TPR (True Positive Rate), FPR (False Positive Rate) and AUC (Area Under Curve) is given, also in comparison with another state-of-the-art technique; achieved results are debated throughout the Section.

A. Dataset Description & SVM Training

Since the proposed algorithm is based on an SVM classification, two distinct image repositories have been taken into account, one for the SVM training phase and the other for the actual testing session. The repository used for the SVM training is the UCID [15] (Uncompressed Colour Image Dataset), consisting of 1338 uncompressed TIFF images of size 512×384 pixels. The test repository is instead the Dresden Image Database [16], consisting of 1488 images whose 736 images of 3039×2014 pixels acquired with a Nikon D70 camera and 752 images of 3900×2616 pixels acquired with a Nikon D200 camera. All the photos are in the NEF vendor private format and have been converted to the TIFF format by using the UFRaw[1] software suite for Linux.

Since the presented algorithm is based on the classification of each tile composing an image under test, the SVM training is performed on a set of $W \times W$ pixel tiles derived from the UCID image dataset. In order to have a consistent set, 40 UCID images have been randomly selected and then split in squared sub-images having the size of $W = 64$ pixels, leading to a total amount of 1920 elements. Future works will concern the reduction of such a search window size. After that, these tiles have been compressed once or twice to build the positive or the negative dataset. The chosen SVM implementation is the MATLAB implementation with the mlp (Multilayer Perceptron) kernel with no autoscaling. A set of quality factors $QF = \{50, 55, ..., 95\}$ consisting of 10 elements is used to perform the training of an equal number of SVMs. Each SVM$_i$, being $i \in \{1, ..., 10\}$, is trained with a positive set composed by the aforementioned 1920 tiles compressed once with QF_i, while the negative set is composed by 9 subsets with the same 1920 tiles compressed twice using QF_j with j≠i, as the first quality factor, and QF_i as the second quality factor. Each SVM classifier can then be used to evaluate each tile of the test image to recognize if it is compressed once or twice; according to the value of the QF of the test image (this is the final QF in case of double compression) the related SVM is selected. Finally, the classification leads to a map of distances from the selected SVM secant hyperplane for each tile composing the test image.

For the test phase, a group of images, taken from the Dresden Image Database, have been spliced with a rectangular patch (randomly selected and aligned with 8×8 grid) covering approximately the 2% of the total surface and two scenarios have been taken into account. In the Single Compressed Patch (SCP) scenario, the patch has been singly compressed with a QF_2 quality factor, while the remaining part of the image is firstly compressed with a QF_1 and then with a QF_2 quality factor. In the Double Compressed Patch (DCP) scenario, that is the dual scenario with respect to the SCP, the patch is double compressed (QF_1 plus QF_2) while the remaining part of the image is single compressed with the quality factor QF_2. The range of the considered quality factors is $QF_i = \{50, 55, ..., 95\}$, as previously introduced.

B. Qualitative Analysis

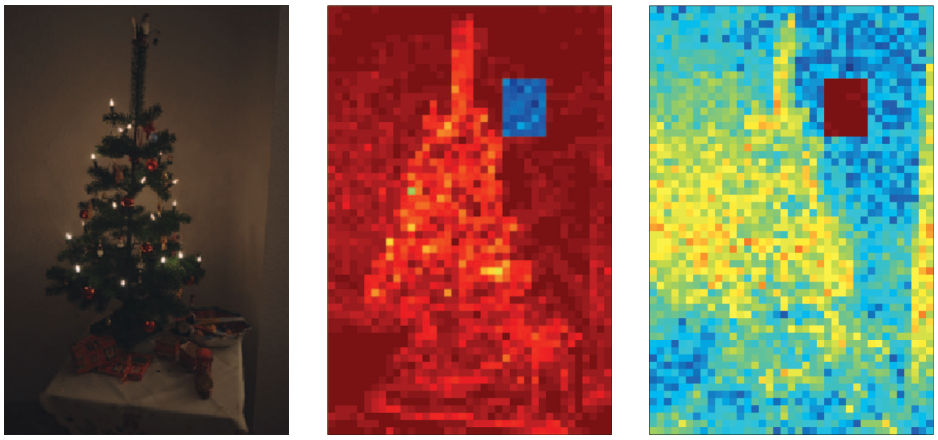
In this subsection, some examples of the achieved results are shown in order to allow an easy access and an intuitive representation of the output of the proposed method.

In Figure IV-B, the heat map obtained for the image named *Christmas Tree* (see Figure 1(a)) for the two test scenarios SCP and DCP (Figures 1(b) and 1(c) respectively) are presented. In particular, in this case, the patch in the top-right region has been compressed once or twice according to the two scenarios with quality factors $QF_1 = 60$ and $QF_2 = 90$.

As the heat map ranges from blue to red, being blue the single and red the double compressed regions, it can be noticed that a clear detection of the single compressed patch in Figure 1(b) is achieved, while in Figure 1(c) some false alarms arise in the central part of the image.

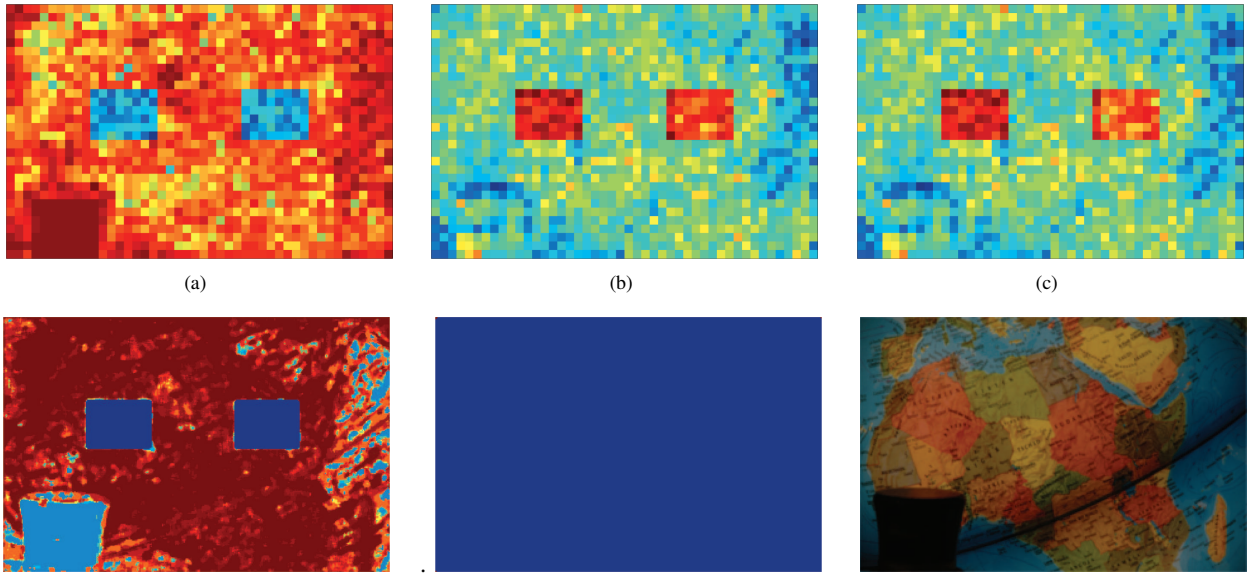
In Figure 2, another example is presented (image named *Globe* Figure 2(f)): two image patches have been taken into account in this circumstance. The splicing attack detection has been performed through the proposed method for the SCP and DCP scenarios (Figures 2(a), 2(b)) with $QF_1 = 60$ and $QF_2 = 90$; both patches are processed in the same manner. For comparison, in Figure 2(d) and 2(e), the results obtained with the method proposed in [2] are pictured (the number of considered DCT coefficients have been set to 15 according to [2]). Even in this double patching setup the detection with the proposed method is clear in both the Figure 2(a) and in the Figure 2(b) case. Instead, while in Figure 2(d) the detection is clear and very detailed, because of fine granularity of the method proposed in [2], in Figure 2(e) the method cannot actually detect any splicing since the algorithm is full-frame based. In fact, it is necessary that a considerable portion of the image was double compressed in order to detect a periodicity in the DCT values histogram. In the DCP scenario, only the spliced regions can contribute to the histogram modification and, since their relative spatial extension is limited, they cannot significantly contribute to the detection. Finally, in Figure 2(c),

[1] http://ufraw.sourceforge.net/



(a) (b) (c)

Fig. 1. Heat maps obtained by the proposed method with $QF_1 = 60$ and $QF_2 = 90$: (a) The *Christmas tree* image from the Dresden Image Database, (b) SCP scenario and (c) DCP scenario.



(a) (b) (c)

(d) (e) (f)

Fig. 2. Heat maps (Two patches splicing): comparison between the proposed method and the method in [2]. Proposed Method: (a) SCP scenario ($QF_1 = 60$, $QF_2 = 90$), (b) DCP scenario ($QF_1 = 60$, $QF_2 = 90$) and (c) DCP scenario ($QF_1 = 60$, $QF_2 = 90$ Left Patch; $QF_1 = 70$, $QF_2 = 90$ Right Patch). Method [2]: SCP scenario ($QF_1 = 60$, $QF_2 = 90$) and DCP scenario ($QF_1 = 60$, $QF_2 = 90$). (f) The *Globe* image from the Dresden Image Database.

the results achieved by the proposed method are presented in the DCP scenario but the two patches have two different initial quality factors ($QF_1 = 60$ left and $QF_1 = 70$ right). The result for the technique in [2] is equal to Figure 2(e) and has not been inserted. It is worthy to notice that the reduced difference between QF_1 and QF_2 in Figure 2(c) leads to less distinctive features for the right spliced region and consequently in a less clear splicing detection.

C. AUCs comparison

In this section a comparison based on the ROC AUCs (Area Under Curve) statistics is presented. Two sets of quality factors, one for the first and one for the second compression are taken: $QF_1 \in \{50, 55, ..., 95\}$ and $QF_2 \in \{50, 55, ..., 95\}$. For each possible couple ($QF_{1,i}$,$QF_{2,j}$) a ROC curve is computed by thresholding at various T the map of distances from the SVM hyperplane; then, for each ROC, an AUC value is computed leading to an $AUC_{i,j}$. The AUC curves presented here are plotted with respect to QF_2 where the $AUC_{i,j}$ values have been averaged over QF_1.

In Figures 3(a) and 3(b) (SCP and DCP scenarios respectively), a comparison between the proposed method (red line) and the algorithm presented in [2] (blue line) is shown. It can

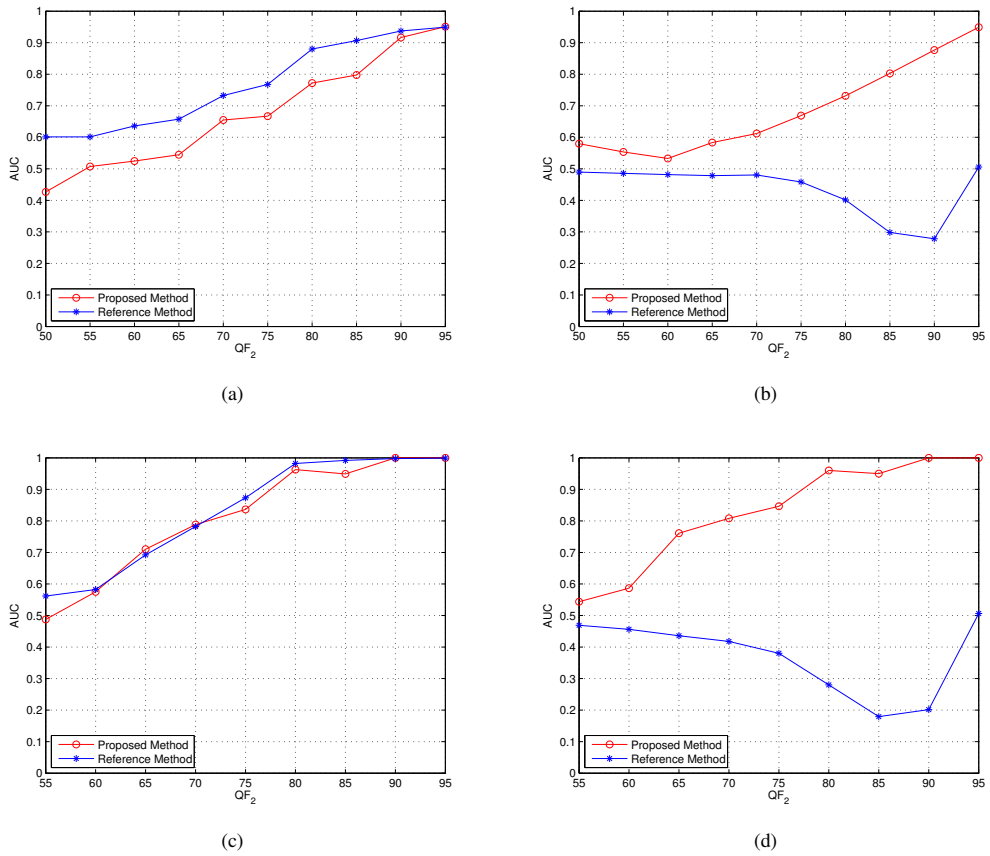


Fig. 3. The AUC comparison. AUC averaged over all the QF_1: (a) SCP and (b) DCP scenarios respectively. AUC averaged over all the $QF_1 < QF_2$: (c) SCP and (d) DCP scenarios respectively.

be appreciated that in the SCP scenario the performances are satisfactory and similar being the red line around 10% lower then the blue line but following a similar trend. In the DCP scenario instead, since locality of the proposed method does not significantly affect the AUC curve trend, its performances remain almost the same while the performances of the method in [2] decrease significantly due to its full-frame dependence. In Figures 3(c) and 3(d), the same comparison is performed, but the AUCs values are averaged only over the values of $QF_1 < QF_2$ in order to exclude the situations represented by $QF_1 \geq QF_2$ where the FPR values are significant as can be seen also in Tables I and II. In fact, it is important to point out that both the techniques present a worst behavior when the the second QF (QF_2) is lower than the first one (QF_1) and this is mainly due to the loss of information caused by the second stronger JPEG compression. On this basis, it can be appreciated in Figure 3(c) that both the methodologies perform better than before and even more similarly.

D. Quantitative Analysis

In this subsection a quantitative analysis is carried on. For each couple ($QF_{1,i}$,$QF_{2,j}$), TPR and FPR values (Tables I and II) are provided. Notice that TPR values are the same for each QF_1 because of the locality of the proposed method since the single compressed region of the image is not affected by the QF_1 variation. Both Table I and II show a low FPR when $QF_2 > QF_1$ and an high FPR value in the dual situation. These data confirms what is known in literature that is when $QF_2 > QF_1$ FPR values tend to 0 while TPR values conversely tend to 100. In both Tables can be noticed that for $QF_2 \geq 70$ TPR values, start to rise. For instance, in both Tables can be noticed that for $QF_2 = 80$ and $QF_1 = 65$ FPRs are 0 and TPRs are 98.65% in the SCP scenario and 95.41% in the DCP scenario.

V. Conclusions

A novel methodology to localize image splicing attack based on first digit features and SVM classifier has been proposed. Given a suspected photo, it can reliably detect if a certain region has been tampered and which are the regions involved. The presented technique shows effectiveness with respect to diverse quality compressions, forgery dimensions and multiple forgeries. Future works will be oriented to strengthen and

TABLE I
FPR & TPR AVERAGED OVER 20 IMAGES OF THE DRESDEN REPOSITORY WITH RESPECT TO QF_2 AND QF_1 - SCP

	FPR (%)										TPR (%)
QF_2/QF_1	50	55	60	65	70	75	80	85	90	95	50 : 5 : 95
50	23.13	24.79	38.60	11.97	18.35	27.16	13.81	19.29	21.95	23.28	16.24
55	21.28	22.81	21.50	4.23	10.41	30.13	13.71	8.64	19.08	22.85	21.40
60	0.23	11.84	16.85	9.46	9.18	7.58	21.50	7.31	17.95	16.74	12.31
65	35.80	45.50	79.68	84.26	86.75	66.09	60.81	91.87	82.20	84.86	83.76
70	0.07	3.24	75.03	84.76	88.85	92.40	69.02	65.88	85.51	89.35	88.85
75	0.02	0.05	9.08	70.86	81.09	86.49	73.50	83.91	90.95	87.02	88.96
80	0.00	0.00	0.01	0.00	2.27	74.04	95.41	98.95	93.25	96.38	98.65
85	0.01	0.00	0.00	0.00	0.00	0.03	63.66	84.58	95.55	75.66	83.92
90	1.02	0.45	0.26	0.24	1.01	0.02	0.40	2.40	99.48	99.53	99.48
95	0.00	0.03	0.01	0.00	0.00	0.01	0.01	0.02	0.07	99.18	99.90

TABLE II
FPR & TPR AVERAGED OVER 20 IMAGES OF THE DRESDEN REPOSITORY WITH RESPECT TO QF_2 AND QF_1 - DCP

	FPR (%)										TPR (%)
QF_2/QF_1	50	55	60	65	70	75	80	85	90	95	50 : 5 : 95
50	16.24	16.25	32.35	7.09	11.34	20.63	7.46	12.29	14.04	16.16	23.13
55	16.75	21.40	21.77	5.89	10.12	29.61	12.77	7.91	17.37	20.74	22.81
60	0.00	9.32	12.31	7.10	6.28	5.32	19.61	5.59	14.14	12.75	16.85
65	29.58	41.30	77.34	83.76	89.89	74.38	69.58	93.09	84.98	84.72	84.26
70	0.00	1.95	60.78	79.90	88.85	95.21	83.23	79.06	91.77	89.79	88.85
75	0.00	0.00	7.15	58.87	80.21	88.96	81.35	75.21	92.71	89.48	86.49
80	0.00	0.00	0.00	0.00	1.77	69.69	98.65	99.79	95.63	98.96	95.41
85	0.00	0.00	0.00	0.00	0.00	0.00	58.99	83.92	96.04	74.29	84.58
90	0.33	0.32	0.39	0.00	0.57	0.00	0.11	0.94	99.48	99.58	99.48
95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	99.90	99.18

consolidate the algorithm in order to cope with a reduced search window size in order also to better address the not-aligned splicing attack. Furthermore, as previously stated, the issue to properly deal with flat and heavily compressed image regions will be studied with the intent to overcome the problem to evaluate a block though its feature descriptor is null.

ACKNOWLEDGMENTS

This work was partially supported by the SECURE! Project, funded by the POR CreO FESR 2007-2013 programme and by the SMARTVINO Project, funded by the PRAF 2012-2015-1.2.e programme, both of the Tuscany Region (Italy).

REFERENCES

[1] L. Zhouchen, H. Junfeng, T. Xiaoou, and T. Chi-Keung, "Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis," *Pattern Recognition*, vol. 42, no. 11, pp. 2492 – 2501, 2009.

[2] T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 3, pp. 1003–1017, 2012.

[3] B. Li, Y. Shi, and J. Huang, "Detecting doubly compressed JPEG images by using mode based first digit features," in *Multimedia Signal Processing, 2008 IEEE 10th Workshop on*, Oct 2008, pp. 730–735.

[4] S. Milani, M. Tagliasacchi, and S. Tubaro, "Discriminating multiple JPEG compression using first digit features." in *ICASSP*. IEEE, 2012, pp. 2253–2256.

[5] E. Kee, J. F. O'Brien, and H. Farid, "Exposing photo manipulation with inconsistent shadows," *ACM Trans. Graph.*, vol. 32, no. 3, pp. 28:1–28:12, Jul. 2013. [Online]. Available: http://doi.acm.org/10.1145/2487228.2487236

[6] I. Amerini, L. Ballan, R. Caldelli, A. D. Bimbo, L. D. Tongo, and G. Serra, "Copy-move forgery detection and localization by means of robust clustering with j-linkage," *Signal Processing: Image Communication*, vol. 28, no. 6, pp. 659 – 669, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0923596513000453

[7] V. Christlein, C. Riess, J. Jordan, C. Riess, and E. Angelopoulou, "An evaluation of popular copy-move forgery detection approaches," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 6, pp. 1841–1854, Dec 2012.

[8] M. Stamm, M. W., and K. Liu, "Information forensics: An overview of the first decade," *Access, IEEE*, vol. 1, pp. 167–200, 2013.

[9] A. C. Popescu and H. Farid, "Statistical tools for digital forensics," in *Information Hiding*, ser. Lecture Notes in Computer Science, J. Fridrich, Ed. Springer Berlin Heidelberg, 2005, vol. 3200, pp. 128–147.

[10] X. Zhao, A. T. S. Ho, and Y. Q. Shi, "Image forensics using generalised Benford's law for accurate detection of unknown JPEG compression in watermarked images," in *Proceedings of the 16th International Conference on Digital Signal Processing*, ser. DSP'09. Piscataway, NJ, USA: IEEE Press, 2009, pp. 518–525. [Online]. Available: http://dl.acm.org/citation.cfm?id=1700307.1700393

[11] X. Feng and G. Dorr, "JPEG recompression detection," pp. 75 410J–75 410J–12, 2010. [Online]. Available: http://dx.doi.org/10.1117/12.838888

[12] Z. He, W. Lu, W. Sun, and J. Huang, "Digital image splicing detection based on markov features in DCT and DWT domain," *Pattern Recognition*, vol. 45, no. 12, pp. 4292 – 4299, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320312002440

[13] E. Acebo and M. Sbert, "Benford's law for natural and synthetic images," in *Proceedings of the First Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging*, ser. Computational Aesthetics'05. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2005, pp. 169–176. [Online]. Available: http://dx.doi.org/10.2312/COMPAESTH/COMPAESTH05/169-176

[14] J. Lukas and J. Fridrich, "Estimation of primary quantization matrix in double compressed JPEG images," in *Proc. of DFRWS*, 2003.

[15] G. Schaefer and M. Stich, "UCID - an uncompressed colour image database," in *Storage and Retrieval Methods and Applications for Multimedia 2004, volume 5307 of Proc. of SPIE*, 2004, pp. 472–480.

[16] T. Gloe and R. Böhme, "The Dresden Image Database for benchmarking digital image forensics," in *Proceedings of the 2010 ACM Symposium on Applied Computing*, ser. SAC '10. New York, NY, USA: ACM, 2010, pp. 1584–1590. [Online]. Available: http://doi.acm.org/10.1145/1774088.1774427