General-Purpose Image Forensics Using Patch Likelihood under Image Statistical Models

Wei Fan, Kai Wang, and François Cayre

GIPSA-lab, CNRS UMR5216, Grenoble INP, 11 rue des Mathématiques, F-38402 St-Martin d'Hères Cedex, France
{wei.fan, kai.wang, francois.cayre}@gipsa-lab.grenoble-inp.fr

***Abstract*—This paper proposes a new, conceptually simple and effective forensic method to address both the generality and the fine-grained tampering localization problems of image forensics. Corresponding to each kind of image operation, a rich GMM (Gaussian Mixture Model) is learned as the image statistical model for small image patches. Thereafter, the binary classification problem, whether a given image block has been previously processed, can be solved by comparing the average patch log-likelihood values calculated on overlapping image patches under different GMMs of original and processed images. With comparisons to a powerful steganalytic feature, experimental results demonstrate the efficiency of the proposed method, for multiple image operations, on whole images and small blocks.**

***Index Terms*—General-purpose image forensics, fine-grained tampering localization, natural image statistics, Gaussian mixture model, patch likelihood**

I. Introduction

During the last a few years, image forensics has become a commonly acknowledged image authentication tool to expose doctored images in a *blind* and *passive* way. Since image forgery creation usually involves various image processing techniques, lots of efforts in the image forensics community have been put to tracing image processing history. Consequently, a large number of forensic algorithms were proposed to reveal different image operations, here to mention a few, *e.g.*, JPEG compression [1], median filtering [2], and resampling [3]. Most of these techniques target at identifying a *specific* image operation, and are indeed powerful at detecting images processed by the image operation under investigation. They in literature are labelled as *targeted* (*ad hoc*) schemes, which however are known to be lack of generality. For example, a powerful forensic feature designed for identifying JPEG compression may not be effective enough to build a median filtering detector. In order to overcome the generality limitation of targeted forensics, *general-purpose* image forensics has recently emerged to cope with multiple image operations.

Though general-purpose image forensics appears to be a more practical (also more challenging-to-design) technique than targeted forensics, its development is still at a very early stage. Probably the only existing work addressing the generality issue of image forensics is the paper of Qiu *et al.*'s [4]. In their forensic strategy, steganography is regarded as a kind of image processing. Therefore, powerful steganalytic features, such as SPAM (Subtractive Pixel Adjacency Matrix) [5], SRM (Spatial-domain Rich Model) [6], and LBP (Local Binary Patterns) [7], are used to build forensic detectors exposing different image operations. In fact, adopting steganalytic features for image forensic purposes is not completely new. Kirchner *et al.* [8] have previously showed the effectiveness of the SPAM feature [5] to serve for median filtering forensics. For JPEG forensic purposes, Li *et al.* [9] also borrowed a feature [10] from steganalysis.

Modern image steganalysis has been developing to successfully detect stego-images with low-rate embedded data by powerful steganography methods. This is realized by the design of complex high-dimensional steganalytic features and the adoption of machine learning methods, such as the ensemble classifier [11] and the SVM (Support Vector Machine) [12]. It is well demonstrated in [4] that these steganalytic features successfully capture the image statistical change caused by different image operations, which actually produce a much higher image modification rate than steganography. The image forensic methods based on steganalytic features [4], [8], [9] as well as a lot of other image forensic algorithms only report, indeed with a very high accuracy, whether a given image of big size has been processed by a certain image operation. However, no results have been provided in [4], [8], [9] concerning the forensic effectiveness of steganalytic features on small image blocks[1], which can be taken as an equivalence to image tampering localization.

Though a simple binary answer may be considered sufficient for steganalysis (linked to Simmons' prisoner problem [13]), it may not be informative enough for image forensics to help people grasp a deep understanding of the image tampering locations and semantics. In this paper, we are particularly interested in exposing the image operation locally in the image. This is due to the common practice of using *cut-and-paste* or *copy-move* to create image forgeries with modified semantics. In this case, it frequently happens that only certain local image areas rather than the whole image have undergone a particular image processing. For revealing the semantics of the tampered image, image forensics should be able to expose image processing on small image blocks.

Based on the above discussion, we would like to emphasize

[1] In this paper, we use two similar words "block" and "patch" with subtle difference. Unless otherwise addressed, "block" is used to referred to a relatively big image area, rectangle but not necessarily square, and "blocks" do not overlap. However, "patch" is usually a small square image area, and "patches" can overlap.

the importance of both the *generality* and the ability to perform *fine-grained tampering localization* for image forensics. With these two goals in mind and following a very different strategy than that of the general-purpose image forensic method [4] inherited from steganalysis, in this paper, we propose a *new*, *conceptually simple* and *effective* way to conduct image forensics based on likelihood comparison under parametric image statistical models. The underlying intuition of the proposed method is reflected by answering the following question: *given an image block, is it more like a natural, original block or a processed one?* In order to answer this question, for each image operation in consideration (or no processing at all), we learn an image statistical model using the GMM (Gaussian Mixture Model) on small image patches. Thereafter, we are able to measure *how likely* a new given image block has been previously processed by a certain image operation, by calculating its average patch log-likelihood value under the corresponding GMM. Meanwhile, the "*naturalness*" of an image block can be measured similarly under the GMM, which is learned on natural, original image patches. Therefore, general-purpose image forensics with good performance of fine-grained tampering localization is thereafter achieved by comparison of the two average patch log-likelihood values. The main advantages of the proposed method over the state-of-the-art image forensics are summarized in the following:

- The forensic methodology of the proposed method is very different from, and conceptually much simpler than the state-of-the-art image forensic methods. It does not need to extract hand-crafted features as almost all current forensic methods do.
- The proposed forensic method is general-purpose, which means it is able to cope with different image operations under the same likelihood comparison framework. It is also easy to be extended when a new image operation is taken into account. The single necessary step is to learn another GMM on image patches, processed by the considered new image operation.
- Since the image statistical models are constructed on small image patches, calculating the average patch log-likelihood value on a small image block is easy. Then, it is straightforward to perform fine-grained image tampering localization using the proposed method.

The remainder of the paper is organized as follows. Sec. II presents the proposed general-purpose image forensic framework, with analysis of the suitability of the GMM for image forensic purposes. Experimental results are reported in Sec. III, with comparisons to a steganalytic feature SPAM [5], on both whole images and small image blocks for fine-grained tampering localization. Finally, we draw the conclusion in Sec. IV.

II. Proposed Method

A. Motivation

As discussed in Sec. I, we hope to devise an image forensic method which is able to perform fine-grained tampering localization. Therefore, rather than analyzing a whole image, we are motivated to work on small image patches. In this paper, we consider the binary classification problem[2] whether a given image block has been processed. To this end, we aspire to find a way to measure the "naturalness" of small image patches. Between the two options "natural" and "processed", if we are able to tell which one is the more likely case for a given image patch, then the classification is straightforward. To do so, it would be ideal if we are able to find a good distribution for natural, unprocessed patches, and that for the processed ones.

As known, the distribution of natural images is very complicated. In the research field of natural image statistics, various attempts have appeared in order to find a good natural image statistical model in the spatial, transformed or filtered domain [14], which is the core of various computer vision problems as well as of image forensics and anti-forensics [15], [16]. Zoran and Weiss [17] firstly proposed to use the GMM for natural image patches. Serving as an excellent image prior model, it is proven to perform very well for various low-level computer vision problems such as denoising, deblurring, inpainting [17], and for JPEG image enhancement [18].

Compared with low-level computer vision problems where the GMM performs excellently as an image prior, image forensics is a very different problem. We are curious about whether the GMM is able to capture the statistical difference between natural, unprocessed images and the processed ones, that the human naked eyes fail to distinguish. To this end, we analyze the GMMs (see Sec. II-B for more details) with 200 components learned on 8×8 image patches extracted from the original images, JPEG images with quality factor 90, and sharpened images with filter parameter 0.5 (see Table I), respectively. If the GMM covariance matrices are able to successfully capture very different image statistics for different image operations, then the GMM will be a good choice for our image forensic task.

For the above mentioned 3 GMMs, the eigenvectors of covariance matrices with size 64×64 of the 4 leading components with the biggest mixing weights are shown in Fig. 1. The first GMM components mainly capture the contrast variability of image patches, whereas the following components reveal the textures and boundaries of objects [19]. However, even in the first 4 components, different patterns can be observed from different GMMs. Compared with Fig. 1-(a)-(d) obtained from the original images, we can see that JPEG compression is well exposed in (e)-(h), where the horizontal and vertical lines shown in the eigenvectors with relatively small eigenvalues actually depict the blocking artifacts present in the JPEG image. Compared with the JPEG images, the eigenvectors obtained from the sharpened images, well reflecting the detail enhancement effect, are even more different from those of the original images.

[2] The image forensic work presented in this paper only considers binary classification, *e.g.*, discriminating between the original and processed image blocks. The generality of image forensics is however embodied by the ability to cope with multiple image operations under the same forensic framework, though respectively considering different binary classification problems. Currently, we do not consider multi-class classification, which we plan to work on in the future.

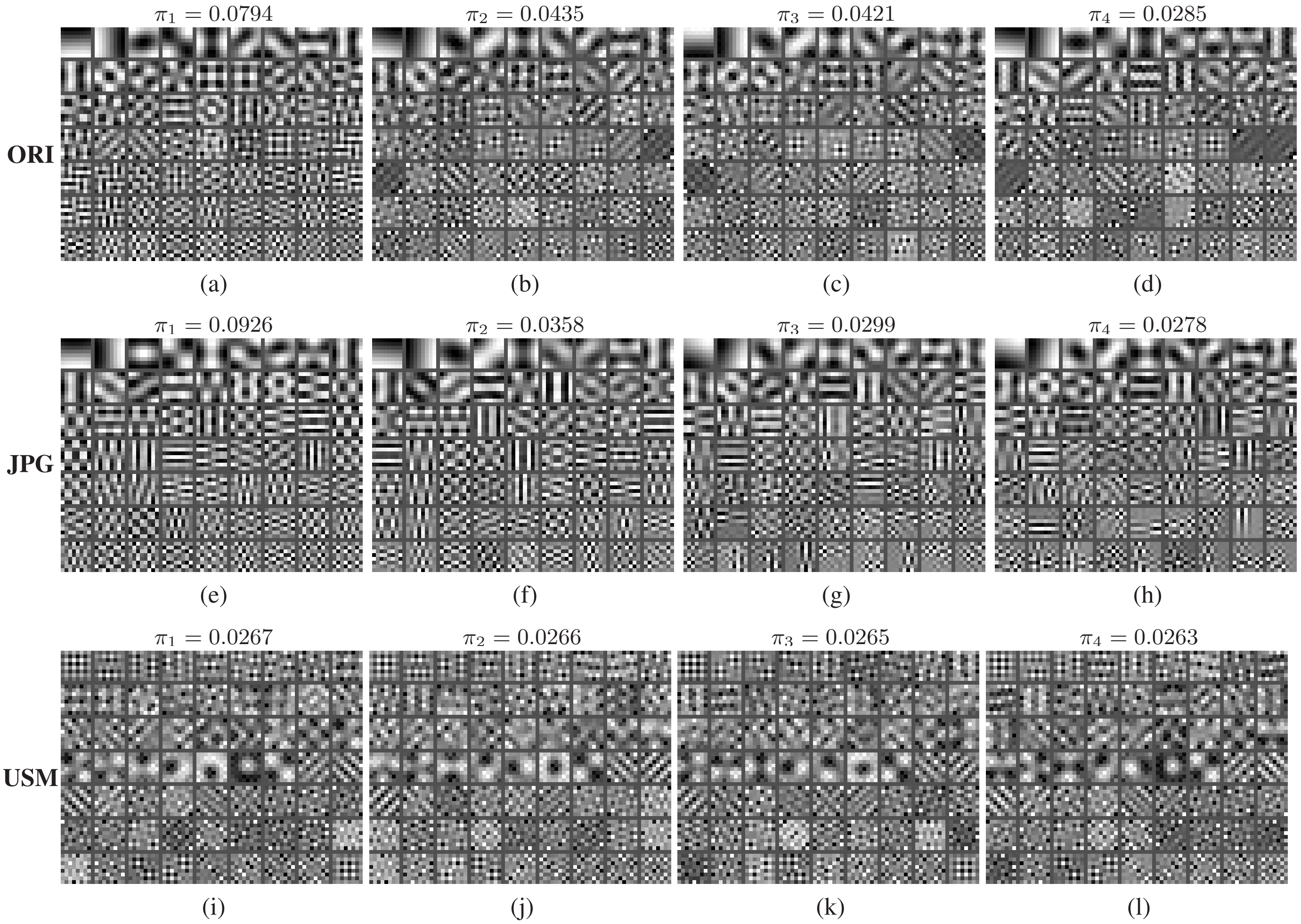

Fig. 1. Eigenvectors of covariance matrices of the first 4 GMM components with the biggest mixing weights. In each subfigure, from the top left to the bottom right, the eigenvectors are sorted according to the descending order of the corresponding eigenvalues. (a)-(d) are obtained from the original images; (e)-(h) are obtained from the JPEG compressed images; and (i)-(l) are obtained from the processed images using the unsharp masking. The GMMs are learned on image dataset GFTR (see Sec. III-A). Different patterns appear in eigenvectors of covariance matrices of GMMs learned on different kinds of images.

Based on the above natural image statistics investigation and experimental analysis of different GMMs learned on different kinds of images, we adopt the GMM as the image statistical model in the proposed forensic method. Under this model, we are able to measure the "naturalness" of a given image patch by likelihood calculation. By further average patch log-likelihood comparison of a given image block, we can thereafter conduct forensic classification. Besides the simplicity and being very expressive for small image patches processed by different operations (as shown in Fig. 1), we also notice that the GMM may also bring generality to the proposed image forensic framework. This is because we can easily integrate a new image processing operation, as long as we are able to correspondingly learn a distinctive GMM.

B. Forensic Analysis Using Image Statistical Models

Based on the discussion and analysis in Sec. II-A, we adopt the GMM to model the distribution of small image patches. Given a generic image patch $\mathbf{x}$ (in vectorized form, its original size is $b \times b$ before stacking), its likelihood under a GMM is computed by:

$$L(\theta|\mathbf{x}) = p(\mathbf{x}|\theta) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \mathbf{C}_k), \tag{1}$$

where π_k, $\boldsymbol{\mu}_k$ and $\mathbf{C}_k$ are respectively the mixing weight, mean and covariance matrix for the k-th ($k = 1, \cdots, K$) GMM component. For sake of brevity, $\theta = \{\pi_k, \boldsymbol{\mu}_k, \mathbf{C}_k | k = 1, \cdots, K\}$ denotes the parameters describing the GMM. The calculated likelihood value $L(\theta|\mathbf{x})$ indicates how likely $\mathbf{x}$ follows the GMM distribution parametrized by θ.

From the patches extracted from images processed by a certain image operation (or no processing at all), the parameters θ of the GMM can be learned using the standard EM (Expectation Maximization) algorithm[3]. Therefore, we build a parametric image statistical model for image patches corresponding to each kind of image operation in consideration.

For the binary classification of our image forensic problem, we are given a generic image (block) $\mathbf{X}$ with size

[3] In practice, we use the unoptimized Matlab code which can be downloaded from: http://www.mathworks.com/matlabcentral/fileexchange/26184-em-algorithm-for-gaussian-mixture-model.

$H \times W$ $(H, W \geq b)$, and we are to give a binary decision about its processing history. From $\mathbf{X}$, we can extract a set of $N = (H - b + 1) \times (W - b + 1)$ overlapping image patches of size $b \times b$: $\{\mathbf{x}_i | i = 1, \cdots, N\}$. Therefore, the average patch log-likelihood (also called expected patch log-likelihood in [17]) of $\mathbf{X}$ can be calculated as $\frac{1}{N}\sum_{i=1}^{N} \log L(\theta|\mathbf{x}_i)$. We assume that $\mathbf{X}$ is original under the null hypothesis $\mathcal{H}_0$, and is processed by a certain image operation under the alternative hypothesis $\mathcal{H}_1$. Under $\mathcal{H}_0$, the patches are assumed to follow the GMM parametrized by θ_0; while under $\mathcal{H}_1$, the patches are assumed to follow the GMM parametrized by θ_1, whose value varies across different image operations. Therefore, the binary classification problem of whether $\mathbf{X}$ has been processed, can be formulated as a simple hypothesis testing problem, and we propose the following test on the difference of two average patch log-likelihood values:

$$\Lambda(\mathbf{X}) = \frac{1}{N}\sum_{i=1}^{N} \log L(\theta_0|\mathbf{x}_i) - \frac{1}{N}\sum_{i=1}^{N} \log L(\theta_1|\mathbf{x}_i) \gtrless \eta. \tag{2}$$

The decision rule of the test is as follows:

$$\begin{cases} \text{reject } \mathcal{H}_0 & \text{if } \Lambda(\mathbf{X}) \leq \eta \\ \text{do not reject } \mathcal{H}_0 & \text{if } \Lambda(\mathbf{X}) > \eta. \end{cases} \tag{3}$$

For image forensic classification, if $\Lambda(\mathbf{X}) \leq \eta$, we classify the image block $\mathbf{X}$ as processed, otherwise as original.

III. Experimental Results

A. Experimental Setting

In this paper, we consider 6 different image operations[4] as summarized in Table I and listed below:

- *Gaussian filtering* is often used for image smoothing, in order to remove noise or to reduce details;
- *JPEG* is one of the most widely used image compression formats today, which is a popular choice of image forgers;
- *Median filtering* is a commonly used image smoothing technique, which is particularly effective for removing impulsive noise. It can also be used to hide artifacts of JPEG compression [20] and resampling [21][5];
- *Resampling* is often involved in creating composite image forgeries, where the size or angle of one source image needs to be adjusted;
- *Unsharp masking* is a popular image sharpening technique, to create less blurry, enhanced image than the original;
- *White Gaussian noise addition*, despite of its rare use in conventional image processing, yet one can find its applications in disguising traces of other image operations, *e.g.*, for JPEG deblocking purposes [20].

For the sake of brevity, the abbreviations in Table I are used to refer to different image operations.

TABLE I
Different image operations considered in this paper.

ORI	no image processing
GF	Gaussian filtering with window size 3×3, and standard deviation 0.5 to generate the filter kernel
JPG	JPEG compression with quality factor 90
MF	median filtering with window size 3×3
RS	resampling with bicubic interpolation to scale the image to 80% of its original size
USM	unsharp masking with window size 3×3, and parameter 0.5 for the Laplacian filter to generate the sharpening filter kernel
WGN	white Gaussian noise addition with standard deviation 2

The natural image datasets used in this paper are created from 545 never resampled, non-compressed TIFF images[6], with various indoor and outdoor scenes. They were taken by 4 cameras of different makes and models, and have been used for image resampling forensics [3] and double JPEG compression forensics [22]. These images are randomly divided into two sets, which respectively include 273 images for training and 272 images for testing. From each original high-resolution TIFF image, we crop 9 adjacent subimages of size 512×512 from its center. Without loss of generality, we only consider grayscale images in this paper. Therefore, we convert the cropped TIFF images into 8-bit grayscale images using the Matlab function `rgb2gray`. In the end, we have $273 \times 9 = 2457$ images in the training dataset GFTR (General-purpose Forensic TRaining), and $272 \times 9 = 2448$ images in the testing dataset GFTE (General-purpose Forensic TEsting).

In order to evaluate the forensic performance of the proposed method on exposing image processing operation on whole images as well as on small image blocks, we consider 3 image (block) sizes $H \times W$: 512×512, 32×32, and 16×16.

B. Forensic Performance

For experimental comparison, we construct forensic detectors based on the well-known 686-dimensional 2nd-order SPAM feature [5], which was initially designed for steganalysis. Its effectiveness in both steganalysis and general-purpose image forensics has been well demonstrated in literature [4], [8]. Moreover, we choose to compare with the steganalytic feature SPAM instead of the SRM (34671-dimensional) [6] and LBP (22153-dimensional) [7] features, partly because of its relatively low dimensionality. We are interested in fine-grained tampering localization and the performance of forensic detectors on small image blocks (down to 16×16 block size with only 256 pixels in our experimental setting). Not only are the feature extraction and detector training of SRM and LBP computationally demanding, but the obtained features from small image blocks may also contain lots of redundancies.

For each image (block) size $H \times W$ and each image operation, a SPAM-based detector is trained using the SVM [12] on the original and the corresponding processed image(s)

[4] Currently, we only consider one parameter setting for each image operation. We will consider multiple parameter settings and include more kinds of image operations in the follow-up work.

[5] Indeed, we consider median filtering or to-be-described white Gaussian noise addition as an image operation motivated by certain image anti-forensic methods. Currently, we however do not test against anti-forensics, which is left as a future effort.

[6] Downloaded from: ftp://firewall.teleco.uvigo.es:27244/DS_01_UTFI.zip and ftp://lesc.dinfo.unifi.it/pub/Public/JPEGloc/dataset/.

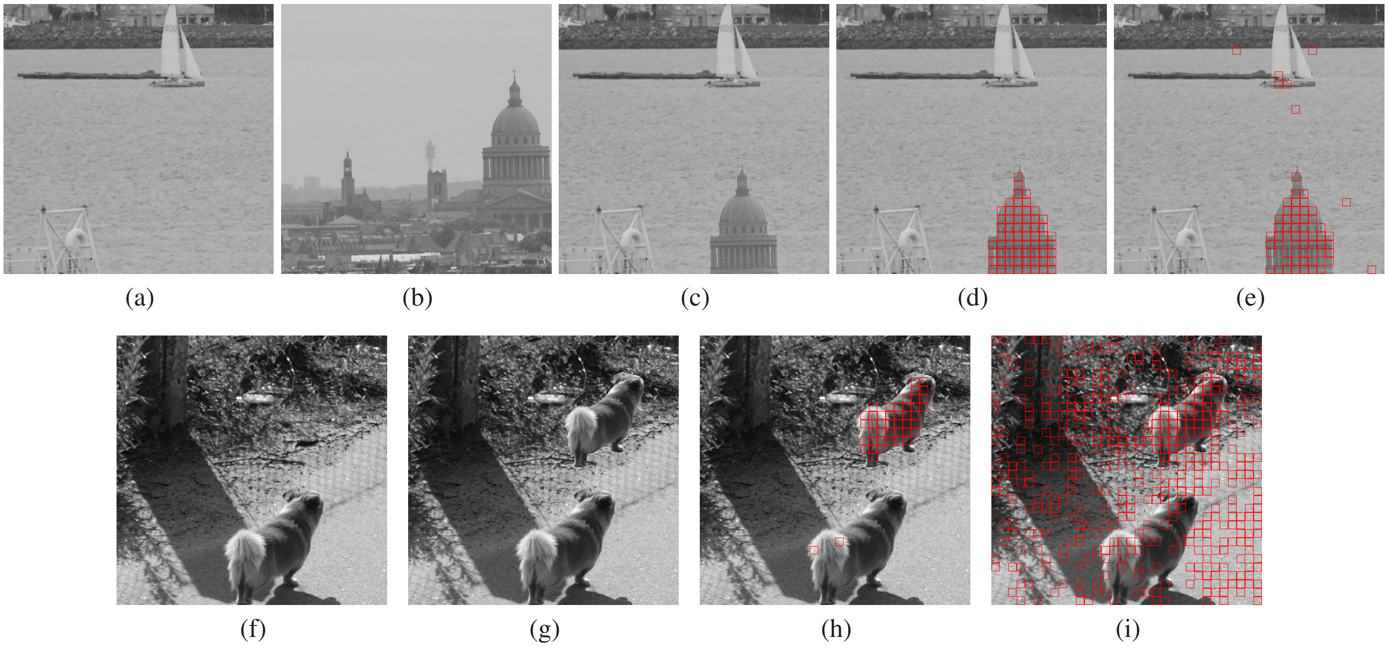

Fig. 2. Example results of fine-grained image tampering localization to reveal forgery semantics, with block size 16×16. Forgery (c) is created by splicing the original image (a) and the church image area extracted from the JPEG image (b) with quality factor 90. (d) points out the JPEG compressed image blocks using the proposed method, and (e) reports the results obtained by the SPAM-based detector [5]. Forgery (g) is created by scaling the dog image area of the original image (f) to 80% of its original size and insert it back to the same image (f). (h) shows the image blocks detected as resampled using the proposed method, whereas (i) presents the results obtained using the SPAM-based detector [5]. Images (a), (f), and the corresponding non-compressed image of (b) are from the GFTE dataset. Results obtained on these two image forgeries indicate the proposed method achieves a better forensic performance on fine-grained tampering localization, because of its high detection accuracy on small 16×16 image blocks.

(blocks) from the GFTR dataset[7]. For learning the GMMs for the proposed method, we consider a patch size of 8×8 (namely $b = 8$), which well balances the richness of the model and the complexity of the learning. We perform the learning procedure on $2457 \times 500 \approx 1.2$ million 8×8 image patches extracted from dataset GFTR, with pixel values scaled to $[0, 1]$ and DC component removed. In the end, we obtain 200 components for each GMM. We have two strategies to set the threshold η in Eq. (3), in order to build the proposed forensic detectors. The straightforward setting is $\eta = 0$, meaning that $\mathbf{X}$ is classified as processed when the test measurement $\Lambda(\mathbf{X})$ indicates that its extracted patches are on average more likely to be under the GMM of the processed image patches. The more sophisticated setting of η can be implemented according to the detector's performance on the GFTR dataset, by maximizing the detection accuracy.

For forensic testing, we use the 2448 images in GFTE dataset for the size 512×512; whereas for the sizes 32×32 and 16×16, 10 blocks are randomly sampled from each (processed) GFTE image, resulting in 24480 blocks corresponding to each kind of image operation.

Table II reports the detection accuracy for different binary classification problems between the original and the processed image(s) (blocks). Here, "SPAM-based" indicates the results achieved by the SPAM-based detectors [5]; whereas "Proposed-S" and "Proposed-T" show results achieved by the proposed detectors by using $\eta = 0$ and the trained η values from the GFTR dataset, respectively. In order to save some computation cost for the proposed method, for the 512×512 image, not all the overlapping $b \times b$ patches are taken into account. Instead, the non-overlapping patches of the image and those of the corresponding image cropped by 4 pixels horizontally and vertically are used. In practice, this simplification strategy does not harm the forensic performance of the proposed method, and can largely reduce the computation cost. We can see from Table II, the proposed detectors are able to achieve comparable performance with

TABLE II
CLASSIFICATION ACCURACY (%) COMPARISON WHEN TESTED ON DIFFERENT SIZES OF IMAGE (BLOCK). THE SPAM-BASED DETECTORS AND THE GMMS ARE TRAINED ON DATASET GFTR. RESULTS ARE OBTAINED BY TESTS ON DATASET GFTE.

		GF	**JPG**	**MF**	**RS**	**USM**	**WGN**
512×512	SPAM-based	99.86	98.20	99.94	96.45	99.73	98.53
	Proposed-S	99.10	97.28	95.69	92.61	99.73	99.45
	Proposed-T	99.82	99.49	99.31	92.67	99.73	99.80
32×32	SPAM-based	99.35	94.18	99.43	89.23	98.76	95.04
	Proposed-S	97.69	95.83	93.81	90.96	99.22	95.50
	Proposed-T	97.73	96.04	93.99	90.96	99.21	97.55
16×16	SPAM-based	98.38	88.00	99.26	78.21	97.82	91.20
	Proposed-S	97.27	94.27	92.88	89.70	98.59	95.58
	Proposed-T	97.37	94.68	93.01	89.72	98.59	95.66

[7] For image block size 32×32 and 16×16, one block is randomly sampled from each GFTR image, for training the SPAM-based detectors. In practice, they have been proven to perform slightly better (around 2% detection accuracy improvement) on image blocks with the corresponding size than SPAM-based detectors trained on 512×512 images.

respect to the SPAM-based ones, when tested on 512×512 images. The proposed method is especially advantageous when tested on small image blocks, especially for the size 16×16. In particular, for resampling, the proposed method obtains 11.5% gain of detection accuracy over the SPAM feature. Even though the proposed method does not outperform the SPAM feature for median filtering, its detection accuracy is still around 93%. Moreover, the forensic performance of the SPAM feature drops, especially for **JPG** and **RS**, when the size of testing image blocks decreases, while ours stays rather stable for all the processing operations in consideration.

The good performance of the proposed method for exposing image processing operation on small image blocks demonstrates its potential for fine-grained image tampering localization. Besides the large-scale test results shown in Table II, here we test on two image forgeries where JPEG compression and resampling are involved. As shown in Fig. 2-(c), it is a composite image forgery by inserting the church image area of the JPEG image shown in -(b) into the original image -(a). The small dog on the upper right of Fig. 2-(g) is actually a copy of the dog image area on the bottom left after resizing it. The two forgeries are divided into non-overlapping 16×16 blocks and are thereafter fed to the forensic detectors for block-wise binary classification. The small red boxes in Fig. 2-(d) and -(h) show the detection results using the proposed method for JPEG compression and resampling, respectively. Compared with results shown in Fig. 2-(e) and -(i) obtained by the SPAM-based detectors [5], we can see that our method successfully reveals, with a higher accuracy, the tampered image areas at a fine-grained level.

IV. Conclusion

Driven by the question whether a given image block is more like a natural, original block or a processed one, this paper adopts the GMM to model the statistics of images processed by different image operations. Given an image block, by comparing the average patch log-likelihood values calculated under two GMMs learned from the original and processed images, we are able to perform binary classification for forensic purposes. Very different from the existing image forensic methods, the proposed method is conceptually simple in methodology, does not need complex feature extraction nor the adoption of the ensemble classifier or the SVM, is easy to be extended for new image operations, and is able to perform fine-grained tampering localization. It is especially advantageous in tracing image processing history on small image blocks, and is able to achieve comparable performance on 512×512 images, compared with the 2nd-order SPAM feature [5].

Future research shall be devoted to investigation of richer image statistical models, multi-class forensic classification, integration of more image operations as well as more of their parameter settings, comparison with detectors based on steganalytic features SRM [6] and LBP [7], and tests against image anti-forensic methods.

Acknowledgment

This work is supported, in part by the AGIR (Alpes Grenoble Innovation Recherche) project TAFFI, in part by the French Rhône-Alpes region through the CMIRA Scholarship Program, in part by the French Eiffel Scholarship under Grant 812587B, and in part by the French ANR Estampille under Grant ANR-10-CORD-019.

References

[1] W. Luo, J. Huang, and G. Qiu, “JPEG error analysis and its applications to digital image forensics,” *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 3, pp. 480–491, 2010.

[2] X. Kang, M. C. Stamm, A. Peng, and K. J. R. Liu, “Robust median filtering forensics using an autoregressive model,” *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 9, pp. 1456–1468, 2013.

[3] D. Vázquez-Padín and F. Pérez-González, “ML estimation of the resampling factor,” in *Proc. IEEE Int. Workshop on Information Forensics and Security*, 2012, pp. 205–210.

[4] X. Qiu, H. Li, W. Luo, and J. Huang, “A universal image forensic strategy based on steganalytic model,” in *Proc. ACM Int. Workshop on Information Hiding and Multimedia Security*, 2014, pp. 165–170.

[5] T. Pevný, P. Bas, and J. Fridrich, “Steganalysis by subtractive pixel adjacency matrix,” *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 215–224, 2010.

[6] J. Fridrich and J. Kodovský, “Rich models for steganalysis of digital images,” *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 868–882, 2012.

[7] Y. Q. Shi, P. Sutthiwan, and L. Chen, “Textural features for steganalysis,” in *Proc. Int. Conf. on Information Hiding*, 2012, pp. 63–77.

[8] M. Kirchner and J. Fridrich, “On detection of median filtering in digital images,” in *Proc. IS&T/SPIE Electronic Imaging*, vol. 7541, 2010, p. 754110.

[9] H. Li, W. Luo, and J. Huang, “Countering anti-JPEG compression forensics,” in *Proc. IEEE Int. Conf. Image Process.*, 2012, pp. 241–244.

[10] C. Chen and Y. Q. Shi, “JPEG image steganalysis utilizing both intrablock and interblock correlations,” in *Proc. IEEE Int. Symp. on Circuits Syst.*, 2008, pp. 3029–3032.

[11] J. Kodovský, J. Fridrich, and V. Holub, “Ensemble classifiers for steganalysis of digital media,” *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 432–444, 2012.

[12] C.-C. Chang and C.-J. Lin, “LIBSVM: A library for support vector machines,” *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 2011, Art. ID 27.

[13] G. J. Simmons, “The prisoners’ problem and the subliminal channel,” in *Advances in Cryptology: Proc. of CRYPTO*, 1983, pp. 51–67.

[14] E. P. Simoncelli and B. A. Olshausen, “Natural image statistics and neural representation,” *Annu. Rev. Neurosci.*, vol. 24, pp. 1193–1216, 2001.

[15] R. Böhme and M. Kirchner, “Counter-forensics: Attacking image forensics,” in *Digital Image Forensics*, H. T. Sencar and N. Memon, Eds. New York, NY, USA: Springer, 2013, pp. 327–366.

[16] W. Fan, K. Wang, F. Cayre, and Z. Xiong, “Median filtered image quality enhancement and anti-forensics via variational deconvolution,” *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 5, pp. 1076–1091, 2015.

[17] D. Zoran and Y. Weiss, “From learning models of natural image patches to whole image restoration,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 479–486.

[18] W. Fan, K. Wang, F. Cayre, and Z. Xiong, “JPEG anti-forensics using non-parametric DCT quantization noise estimation and natural image statistics,” in *Proc. ACM Int. Workshop on Information Hiding and Multimedia Security*, 2013, pp. 117–122.

[19] D. Zoran and Y. Weiss, “Natural images, Gaussian mixtures and dead leaves,” in *Adv. Neural Inf. Process. Syst.*, 2012, pp. 1736–1744.

[20] M. C. Stamm and K. J. R. Liu, “Anti-forensics of digital image compression,” *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 1050–1065, 2011.

[21] M. Kirchner and R. Röhme, “Hiding traces of resampling in digital images,” *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 4, pp. 582–592, 2008.

[22] T. Bianchi and A. Piva, “Image forgery localization via block-grained analysis of JPEG artifacts,” *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 1003–1017, 2012.