IET Image Processing

Review Article

# Bibliography of digital image anti-forensics and anti-anti-forensics techniques

ISSN 1751-9659
Received on 1st December 2018
Revised 14th April 2019
Accepted on 2nd June 2019
E-First on 1st August 2019
doi: 10.1049/iet-ipr.2018.6587
www.ietdl.org

*Muhammad Ali Qureshi*[1], *El-Sayed M El-Alfy*[2] ✉
[1]*The Islamia University of Bahawalpur (IUB), 63100 Pakistan*
[2]*King Fahd University of Petroleum and Minerals (KFUPM), Dhahran 31261, Saudi Arabia*
✉ *E-mail: alfy@kfupm.edu.sa*

**Abstract:** With the massive increase of online content, widespread of social media, the popularity of smartphones, and rise of security breaches, image forensics has attracted a lot of attention in the past two decades alongside the advancements in digital imaging and processing software. The goal is to be able to verify authenticity, ownership, and copyright of an image and detect changes to the original image. However, more sophisticated image manipulation software tools can use subtle anti-forensics techniques (AFTs) to complicate and hinder detection. This leads security professionals and digital investigators to develop more robust forensics tools and counter solutions to defeat adversarial anti-forensics and win the race. This survey study presents a comprehensive systematic overview of various anti-forensics and anti-AFTs that are proposed in the literature for digital image forensics. These techniques are thoroughly analysed based on various important characteristics and grouped into broad categories. This study also presents a bibliographic analysis of the-state-of-the-art publications in various venues. It assists junior researchers in multimedia security and related fields to understand the significance of existing techniques, research trends, and future directions.

## 1 Introduction

Over the past years, several advancements have been witnessed in multimedia, communication, mobile, and storage systems technologies. This has led to a massive increase in online sharing of digital images over social and professional networks such as Instagram, Flickr, Pinterest, and Google Photos. For instance, Table 1 demonstrates the number of unique monthly visitors of top 15 popular online image-sharing websites as of May 2018 [http://www.ebizmba.com/articles/photo-sharing-sites]. Moreover, there are many other generic file sharing systems such as Google Drive, OneDrive, Dropbox etc. Images were often considered as a reliable visual source for recording and authenticating important events and activities from personal to official levels. However, with the availability of modern image editing software tools, such as Adobe Photoshop, Corel Draw, and Pixlr, images become prone to more manipulation than ever before. Various ethical or unethical reasons can drive the changes. This makes the trustworthiness of digital images a major concern. At certain circumstances, these changes are destructive and it is crucial to detect their presence and locate the altered regions in the image [1].

The history of image forensics also covers a lot of interesting examples where the digital photos of politicians, national heroes, or celebrities were manipulated for ethical/unethical reasons. Fig. 1 shows an example of a forged portrait of the former US President Abraham Lincoln in 1860 that could be considered as the first Photoshopped image. In this photo, Lincoln's head was stitched with the body of a Southern politician, John Calhoun. The intention was to portray the US President like an American hero and at that time no heroic style photo of Lincoln was captured. In this example, the original photo is shown on the right side whereas the tampered one is on the left side. Another example of a Photoshopped image is shown in Fig. 2, where Queen Elizabeth Bowes-Lyon (mother of Queen Elizabeth II) and the Canadian Prime Minister, William Lyon Mackenzie, in Banff, Alberta are

**Table 1** Examples of most popular online image-sharing websites demonstrating the estimated number of unique monthly visitors

| Image-sharing sites | Monthy unique visitors (in million) |
|---|---|
| instagram | 100 |
| imgur | 87.5 |
| flickr | 80 |
| photobucket | 60 |
| deviantArt | 45 |
| shutterfly | 20 |
| tinyPic | 9 |
| weHeartIt | 8.1 |
| imageShack | 8 |
| imageVenue | 5.75 |
| imageBam | 5.5 |
| smugMug | 5 |
| picasa | 4.5 |
| twitpic | 4 |
| snapfish | 3.4 |

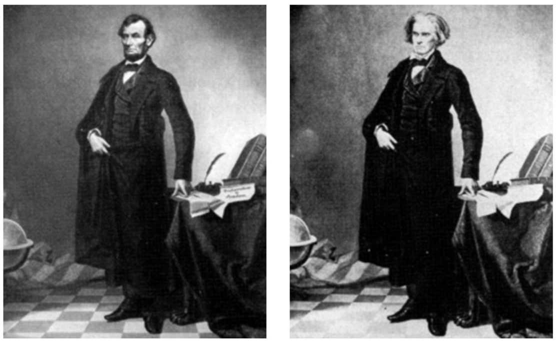

**Fig. 1** *1860: US President Abraham Lincoln Photoshopped portrait (right image), and source image of Southern politician John Calhoun whose body was stitched with Lincoln's head (left image)*

shown together without King George VI in the original image. The tampering was made during the election campaign of the King to present himself in a more powerful position as standing alone with the Queen. There are several other examples of forged images over the past years. Nowadays, the quote that ‘seeing is believing’ is undoubtedly sceptical [2]. For more examples of image tampering and an overview of the historical development of image forensic, there are a number of surveys and overview papers in the literature [1, 3–7].

A variety of algorithms have been proposed for image forensic investigation of digital images. They aim at detecting image manipulation/doctoring such as image enhancement (contrast adjustment, histogram equalisation (HE), colour modification), content alteration (e.g. copy-move, copy-paste, splicing, seam carving), and geometric changes (e.g. rotation, scaling, cropping) [4, 8–14]. These algorithms are classified into various categories such as active or passive, blind or semi-blind, forgery dependent or forgery independent, and pixel-based, format-based, camera-based, physics-based or geometric-based [9, 15–18].

Active methods involve adding extra information to images as a preventive measure, which can be checked at later stages to confirm the authenticity of digital images. The examples of active methods are fragile and semi-fragile watermarks and digital signatures. These methods are intrusive and may drastically affect the content quality. They are even not suitable for images that have been already uploaded on the Internet without adding this preventive information. Passive methods are therefore used to overcome the limitations of active methods. Fig. 3 shows a tree diagram listing the different classes of passive digital image security aspects discussed in the literature. The tampering detection techniques are divided into four broad categories: (i) image copy-move forgery detection, where the source and destination for tampering image regions belong to the same image (see Fig. 4), (ii) In image splicing, the tampered regions belong to another image (see Fig. 5) [19], (iii) Image retouching is also a subtype of tampering where the details of the image are enhanced to make it more prominent, especially used by professionals to improve the quality of photographs or as a form of arts, and (iv) Finally, image inpainting which is also considered as copy-move forgery except for the difference that the source regions belong to different locations of the image [20, 21]. It is usually used to restore the missing pixels, cracks in an image or to remove some objects.

Image tampering processes somehow leave some traces of manipulation in the tampered image, though these traces may be invisible to naked human eyes when the original image is not available. Image forgers may apply various techniques and tricks to hide the tampering traces making detection a bit harder or impossible by traditional forensic techniques [22–26]. Therefore, it is becoming extremely critical to consider these anti-forensics techniques (AFTs) in developing more robust image forgery detection methods. Although a lot of research has been carried in developing more robust image tampering detection techniques, little effort has been made into overcoming intelligent counterfeiters who deliberately hide the tampering traces during the forgery process to fool the detector. The AFTs, however, aim to deceive image forensic by hiding tampering fingerprints. On the other hand, to rebuild the credibility of forensics, counter counter-forensics or anti-anti-forensics is also being used.

In the literature, various anti-forensics and anti-anti-forensic algorithms have been proposed to hide image manipulations, fingerprints, and countering the hidden process, respectively. After deep consultation of literature, we have found that there is no survey on image forensics so far that provides a detailed discussion on anti-forensics and anti-anti-forensics methods. The scope and main contributions of this paper are: (i) survey the state-of-the-art of AFTs and Anti AFTs (AAFTs), (ii) provide readers with a comprehensive review and analysis of 81 most relevant publications, and (iii) present a summary of main differences of available methods. The analysis of the covered publications, collected from Scopus.com and Google Scholar, has revealed that 77 of them address image anti-forensic techniques and 27 of them are concerned with image AAFTs. Fig. 6 shows the publication trend and distribution into conference proceedings and journal articles. As demonstrated in this figure the majority of publications focus on image AFTs, yet in recent years the trend of publications on AFTs is declining whereas more publications appear on AAFTs.

The rest of the paper is organised as follows. In Section 2, we discuss anti-forensic methods, while AAFTs are discussed in Section 3. Finally, we conclude the paper with some critical thoughts and future research directions in Section 4.

## 2 Anti-forensics techniques (AFTs)

Based on the employed methodology, image anti-forensic techniques are grouped into different categories and discussed in detail in the following subsections.

### *2.1 Contrast enhancement (CE)*

In cut-and-paste (or splicing) image forgery, the pasted regions and the background of the original image usually have variations in the contrast level due to different lighting conditions between the two images. To hide this visual difference, forged images are typically post-processed using CE operators to adjust the global contrast and brightness. In order to detect retouched images by CE, the image forensic analysis depends on the assumption that the histogram of a natural image is generally smooth whereas the histogram of a contrast-enhanced image has a wide envelope with peak and gap artefacts [27]. Thus, the CE detectors utilise features derived from the first-order statistics such as peaks and gaps analysis of image histograms [27–30]. To fool histogram-based CE forensic detectors, different techniques have been proposed as explained in the following subsections.

*2.1.1 Trace hiding:* The methods in this category hide the peak-gap artefacts in image histogram to be undetectable by blind image forgery detectors. They use variants of CE operators by performing local random dithering operations with the addition of Gaussian noise. Cao *et al.* [31] performed local random dithering operations in the primary mapping of CE to hide the peak-gap artefacts in the pixel grey-level histogram of the contrast-enhanced image. The same authors in [32] also proposed trace forging and trace hiding

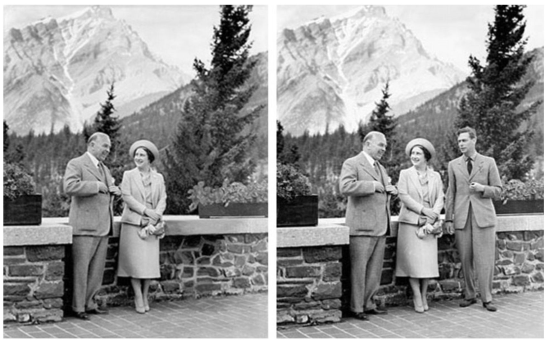

**Fig. 2** *1939: An example of doctored photo of Queen Elizabeth (right image) and original version with King George VI (left image)*

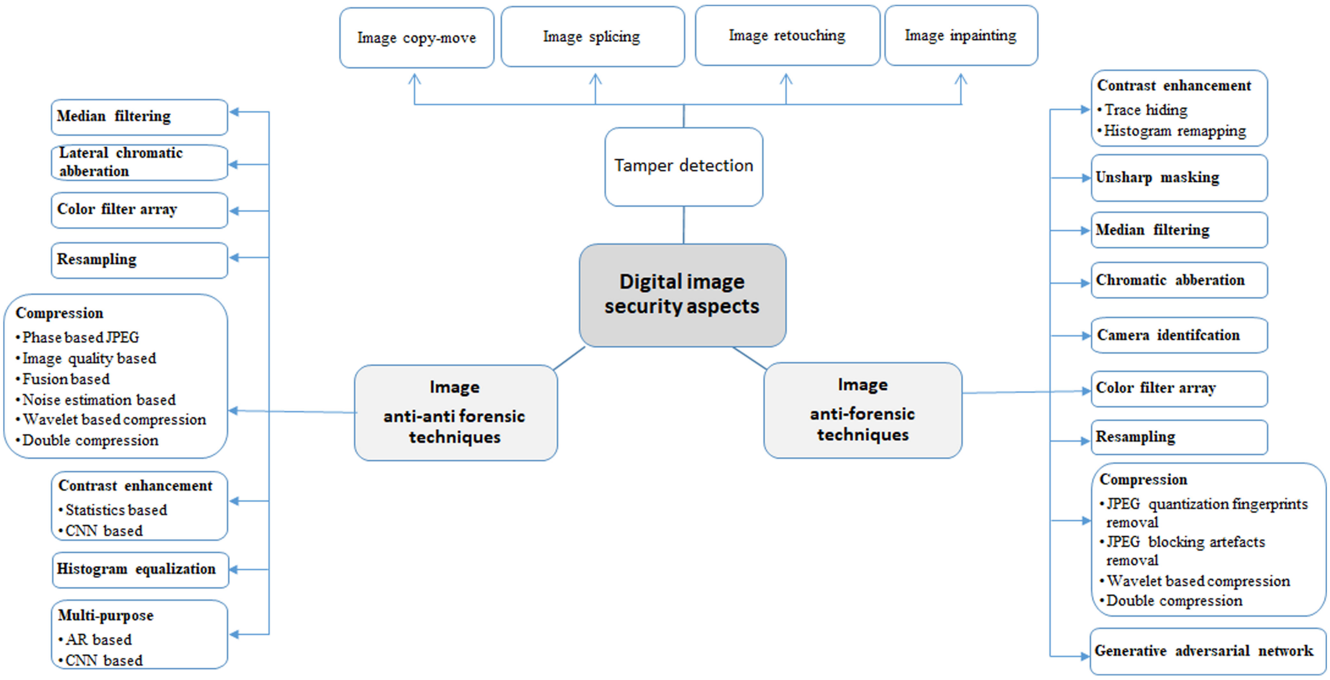

**Fig. 3** *Framework for grouping the different digital image security techniques*

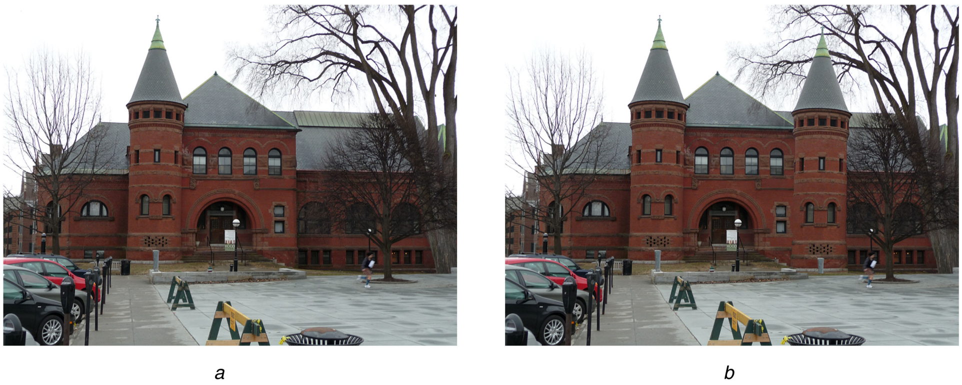

**Fig. 4** *Example of copy-move forgery*
***(a)*** Original image, ***(b)*** Tampered image

attacks against CE detectors. The results showed significant reductions in detection accuracy compared to [27, 28]. Kwok *et al.* [33] used Internal Bit Depth Increase method for further improvement in concealing peak and gap artefacts with good visual image quality.

*2.1.2 Histogram remapping:* In these methods, the histogram of the contrast-enhanced image is modified according to that of the original image using a pixel mapping technique. In this regard, Barni *et al.* [34, 35] proposed a universal anti-forensic technique to counter histogram-based manipulation detectors. A general optimal framework was proposed by Comesana Alfaro and Pérez González [36] in which the authors derived the optimal changes to the histogram to defeat existing histogram-based forensic detectors. Since CE detectors based on first-order statistics are shown to be less robust by the above-mentioned anti-forensic techniques [37], CE detectors based on other metrics were proposed. Lin *et al.* [30] used the inter-channel dependency due to colour image interpolation to detect CE in an image. Another recent algorithm proposed by De Rosa *et al.* [37] used the Grey Level Co-occurrence Matrix (GLCM) of an image to determine whether it is CE manipulated or not. The use of second-order statistics of the image is shown to be more effective against anti-forensic techniques targeted at histogram-based detectors. The performance

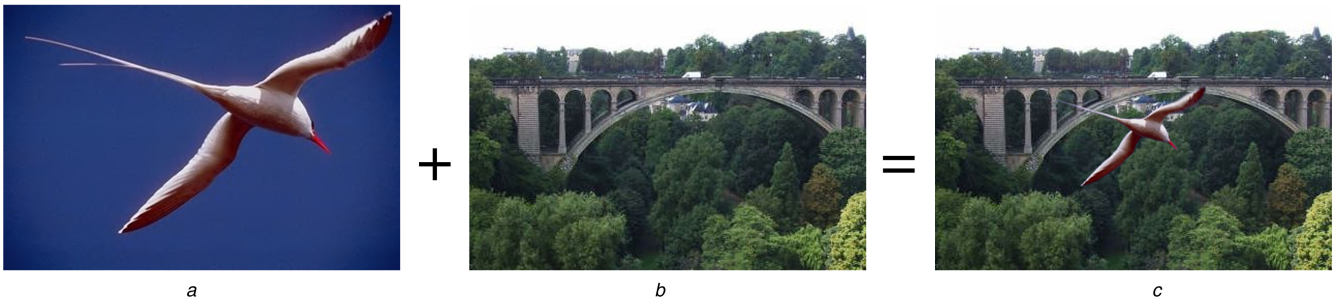

**Fig. 5** *Example of image splicing*
***(a)*** and ***(b)*** original images, ***(c)*** Tampered image created by combining (a) and (b) [19]

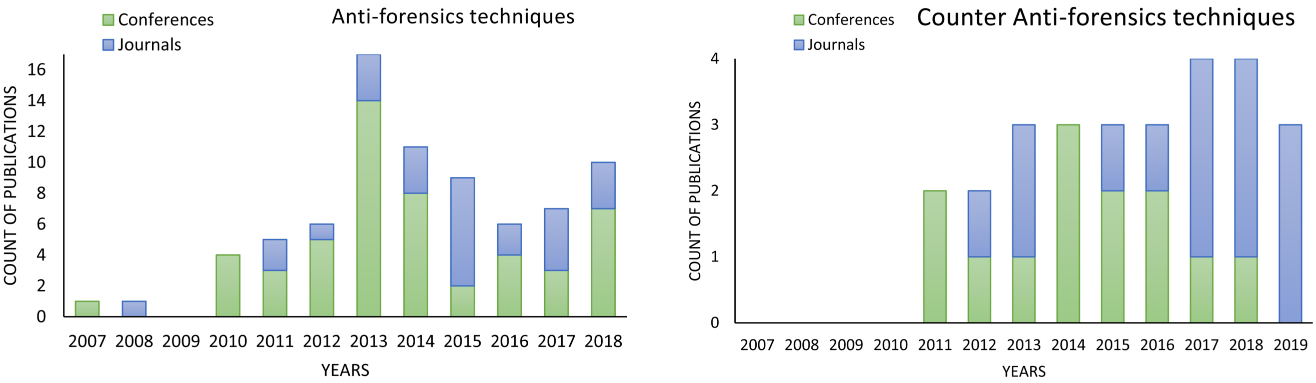

**Fig. 6** *Publication trend and distribution into conference proceedings and journal articles*

evaluation was carried using a large database and results were compared with the first-order [27] and second-order [37] statistics-based CE detectors. It is, therefore, suggested to use a third or even higher-order statistics to develop more robust and efficient CE-based AFTs. Ravi *et al.* [38] proposed an effective AFT for CE using a variation of the total variation (TV) norm optimisation. The algorithm performed CE on an image by solving an optimisation problem. The problem was formulated such that even though the solution looks contrast enhanced with high visual quality, it has lower TV and characteristics similar to the original image.

### *2.2 Unsharp masking*

Unsharp masking (USM) is a type of sharpening that is commonly used in the final stages of image retouching to disguise the blurring alterations in an image. It is available in popular image editing tools such as Photoshop and GIMP. However, it can be detected using techniques such as those proposed in [39, 40]. Laijie *et al.* [41] proposed an AFT for USM sharpening detection by adding random local dithering to remove the overshoot artefacts in image edges and the abrupt change in the histogram. The authors demonstrated the effectiveness of the technique with experimental work carried out on a large image database with different parameter settings and without degradation in the image visual quality.

### *2.3 Median filtering*

Median filtering (MF) is another preprocessing step involved in image forensics to remove the noise with significant changes in the image details. Kang *et al.* [42] proposed a robust method for detecting median filtering based on an autoregressive (AR) model. To counter median filtering detectors, Nguyen *et al.* [43] explored the characteristics of median filtering operation based on the local features derived from image blocks in terms of median distribution, the occurrences of block centre grey level, the quantity of grey levels, distribution of centre grey level in sorted grey levels, and the first occurrence of block centre grey level in sorted grey levels. The fingerprints of median filtering are concealed in uncompressed images by modifying random pixels based on the distribution of block median and occurrence of block centre grey level. Smooth blocks are expected to have block centre value occurring more frequently and high probability of having more than one median value within a block. So the blocks with high values for the distribution of block median and occurrence of block centre grey level are discarded and a random small perturbation is added to those blocks amongst the rest which have standard deviation higher than the threshold.

Wu *et al.* [44] proposed an AFT for median filtering by modelling the distribution of pixel differences as a Generalized Gaussian Distribution (GGD). The GGD parameters were estimated using simple regression. Similarly, Fan *et al.* [45] proposed an optimisation framework considering the distribution of pixel differences as a prior term for median filtering anti-forensics and image reconstruction. Nguyen *et al.* [43] proposed an AFT for MF by adding noise to highly textured blocks. Fontani and Barni [46] proposed an optimisation framework for anti-forensics of median filtering using linear convolution filters with size $3 \times 3$ while maintaining the image quality in terms of Peak Signal-to-Noise Ratio (PSNR) while reducing the cost of features that detect non-linear filtering.

The algorithms discussed in [43–46] performed well for a specific MF detection technique. A comprehensive performance evaluation of different MF detectors [47–49] against anti-forensic attacks was provided by Sharma *et al.* [50]. The technique is different from [45] in the sense that the work in [45] hides the MF traces in a postprocessing step without using the original image information whereas in [50] both original and median filtered image information are used to hide the traces of MF in the resultant image. A recent approach based on the generative adversarial network (GAN) is proposed in [51] to deceive existing median filtering techniques.

### *2.4 Anti-forensics for chromatic aberrations*

The Lateral Chromatic Aberration (LCA) is an imaging trace introduced in optical imaging systems due to the inability of the camera lens to focus all wavelengths of light at the same point. It results in relative contraction or expansion between colour channels of an image and produces imperceptible colour fringes around image edges. In earlier work, the localised inconsistencies in LCA have already been used to detect image splicing in digital images [52–55]. To hide the traces of LCA presenting in a tampered image due to splicing, Mayer *et al.* [56] proposed an anti-forensic technique by calculating the LCA locally at an image location. The deviations from the estimated chromatic aberration due to tampering was removed from the tampered image. The results showed the effectiveness of the technique in hiding tampering traces in digital images.

### 2.5 Countering camera identification

Source identification based image forensics are also being used. Among these, Photo-Response Non-Uniformity (PRNU) sensor noise pattern-based source detection is efficient due to the uniqueness of PRNU noise pattern for every digital camera. Sengupta *et al.* [57] used median filtering to suppress the PRNU noise to fool the PRNU sensor noise-based source identification. The results showed the effectiveness of the scheme in terms of source anonymity using inverse of Peak-to-Correlation Energy ratio measure compared to the state-of-the-art. Similarly, Villalba *et al.* [58] proposed a two-stage algorithm to counter the forensic techniques and forge the image source based on sensor noise and wavelet transform such that it points to an unrelated device without requiring physical access to the source camera or its related information. Thus, this algorithm provided a more realistic approach to destruct the identifiable data and falsify the image camera identification.

Deep learning has recently attracted much attention, especially in computer vision applications. It has been found effective to design a camera identification model based on employing a Convolutional Neural Network (CNN) [59]. However, Chen *et al.* [60] have recently proposed a new AFT based on a GAN to fool the CNN-based camera identification model at the cost of visually imperceptible distortion introduced in the tampered images. The results showed 98% accuracy in fooling CNN-based camera model identification classifier.

### 2.6 Anti-forensics of colour filter array (CFA) pattern

Most of today's digital cameras contain only one type of colour sensor at each pixel. Colour interpolation algorithms are then used to produce high-quality images. Identification of these colour interpolation algorithms is also used in digital image forensics. The CFA interpolation introduces periodic local correlation pattern between neighbouring pixels and can be used to find the traces of demosaicing (or colour restoration) [61, 62]. To hide the traces of these periodic local correlation, Böhme and Kirchner [63] proposed an AFT to alter the raw image pixels by introducing CFA patterns in the tampered images. The method was based on synthesising a linear dependency among image pixels while minimising the overall distortion by searching a pre-filter to estimate raw samples that are captured by the camera sensor. Finally, the entire image satisfying the linear dependency was reconstructed by estimating the raw samples using a bilinear interpolation kernel. The technique was validated on a large image database and compared with state-of-the-art techniques. The results showed the effectiveness of the CFA synthesis in hiding tampering traces without loss of image visual quality. Guangling *et al.* [64] proposed an image AFT based on CFA structure detection using dictionary re-demosaicing. A sparse representation-based dictionary was constructed from the tampered image. An arbitrary CFA demosaicing was performed to re-demosaic each dictionary atom. Finally, the tampered image without CFA traces was reconstructed using the re-demosaiced dictionary. The experiments showed the effectiveness of the technique in terms of visual quality and strong CFA characteristics. Similarly, Chuang and Wu [65] also proposed two anti-forensic techniques to conceal the colour interpolation traces. These techniques used parameter perturbation and algorithm mixing to prevent identification. The results showed excellent performance to hide the CFA traces and fool the CFA identification based image forensics without loss of image quality.

### 2.7 Anti-forensics resampling

Image resampling is related to geometric transformation. In image forensics, to make the tampering more realistic, the forged regions within the tampered image may undergo geometric transformations like scaling, rotation etc. During this phase, resampling is conducted followed by the interpolation. With regular sampling, the interpolation weight sequences are periodical, and thus they can be used to detect the resampling operation. In the literature, various techniques exist for resampling detection [66–69]. To hide the traces of resampling, different approaches have been proposed [70, 71]. Kirchner and Böhme [71], proposed an efficient AFT to hide the traces of periodic artefacts left due to the interpolation operation. The results showed good performance against the detection of resampling proposed in [66–69].

### 2.8 Anti-forensics for compression

Images are often compressed for effective storage and transmission. Among the compression techniques that are most popularly used are Joint Photographic Experts Group (JPEG) and JPEG 2000, which use discrete cosine transform (DCT) and discrete wavelet transform (DWT)-based transformation, respectively. To detect this type of image manipulation, image forensic tools benefit from the fact it is lossy compression. Coding fingerprints due to the compression are widely used in image forensics to detect the tampering operation. Single and double JPEG compression detection [72, 73] or other classes of image compression including the JPEG image forensics based on DCT/DWT coefficient quantisation fingerprints [74], blocking artefacts [75], and wavelet-based techniques [76], are important digital image forensic techniques used for detecting the presence of tampering in digital images. To fool these tools, several approaches have been proposed to remove these traces. A summary of several methods is shown in Table 2.

For instance, detectors can be tricked by mostly relative anti-forensic techniques by dithering the DCT coefficients for imitating the original uncompressed histogram followed by filtering to remove blocking artefacts, or adding some noise to erase the compression history. Fig. 7 shows a framework for the anti-forensic operation in JPEG image compression. Stamm *et al.* [97] proposed AFT to suppress the JPEG blocking artefacts.

In lossy JPEG compression, quantisation and blocking artefacts are commonly observed. The quantisation artefacts are produced in quantisation stage where DCT coefficients are made closer to the multiples of the step size and produce irregularities in DCT coefficients. This quantisation artefact is clearly visible in the histogram of DCT coefficients. This artefact is also known as transform coefficient quantisation artefact. The blocking artefacts introduced due to pixel value discontinuities across block boundaries during block-based DCT computation. These compression artefacts are exploited in AFTs by hiding their traces and are discussed in the following subsections.

*2.8.1 AFTs of JPEG quantisation fingerprints removal:* Recently, disguising the DCT coefficients quantisation fingerprints produced by JPEG compression were extensively cast-off and studied by researchers in AFTs. Targeting to hide DCT quantisation artefacts, Stamm *et al.* [77, 80] proposed the addition of dither noise $d$ to DCT coefficients $z$ in each subband at the cost of some distortion being introduced within the acceptable limits.

$$z_{\mathrm{d}} = z + d \tag{1}$$

where $z_{\mathrm{d}}$ represents the dithered noise and it approximates the distribution of the unquantised coefficients(e.g. AC coefficients follow Laplacian distribution and DC coefficients follow uniform distribution). The additive noise distribution is chosen in such manner that the DCT coefficients are not clustered around integer multiples of the step size. In [77], quantisation gaps in the DCT histogram are diminished by expanding the coefficients with an additive noise also called anti-forensic dither. The additive noise distribution mainly depends on DCT coefficients, the quantisation step, and the estimation of DCT coefficients distribution before compression. As a result, the DCT coefficient distribution in the tampered image becomes closer to the distribution of the unquantised coefficient. In [80], the distribution of the image's transform coefficients was estimated prior to the compression. The

**Table 2** Summary of JPEG AFTs in chronological order

| Reference | Year | Target | Experimental setting | Remarks |
|---|---|---|---|---|
| Stamm *et al.* [77] | 2010 | quantisation | UCID [78] (244 greyscale images, JPEG compression with quality factor (QF): 90, 70, 50) | add anti-forensic dither to hide quantisation artefacts, success rate: 95.9% (90 QF), 92.6%(70 QF), 81.56% (50 QF) |
| Stamm and Liu [79] | 2010 | quantisation | UCID [78] (244 greyscale images, SPIHT at 2bpp (bits/pixel)) | add anti-forensic dither to the wavelet coefficients of highly textured blocks of the compressed image; 100% success rate |
| Stamm and Liu [80] | 2011 | Quantisation & Blocking | UCID [78] (1338 greyscale images, 1st experiment: JPEG compression with QF: 90, 70, 50; 2nd experiment: SPIHT at 2bpp) | Add anti-forensic dither to hide quantisation fingerprints followed by deblocking to remove traces of blocking artefacts, able to defeat existing compression and history-based forensic detectors (e.g. JPEG and SPIHT compression fingerprints, double compression, image origin falsification, and cut-and-paste image forgery); success rate: 100% for JPEG and 98.2% for SPIHT |
| Sutthiwan and Shi [81] | 2011 | Quantisation | UCID [78] (1138 for training, 200 for testing, 20 runs, (QF1, QF2)∈ {50, 60, 70, 80, 90}) | Based on Shrink and Zoom (SAZ), mislead double JPEG compression detectors, scaling factor = 0.9 to keep a comprise between image quality and true positive rates after attack, close to 100% success in most cases |
| Jiang *et al.* [82] | 2013 | quantisation & blocking | UCID [78](1338 greyscale images, QF ∈ {55,65, 75, 85, 95}) | eliminate blocking artefacts with low computational complexity by predicting the added noise, detection accuracy over $99\%$ |
| Fan *et al.* [83] | 2013 | Blocking | UCID [78] (1338 greyscale images of size $512 \times 384$, QF ∈ [30, 84] | removes traces of JPEG compression in both spatial and transform domain, uses a two-round TV-based deblocking and histogram smoothing operation, better tradeoff between image quality and hiding compression fingerprints, PSNR improvement by 4.2 dB |
| Barni *et al.* [85] | 2014 | quantisation | created 64 histograms of DCT coefficients for each image of 2000 greyscale uncompressed images and compressed with multiple QF, 25 testing images from multiple sources | universal AFT capable of misleading multiple-JPEG compression detectors. DCT sub-band histograms of a multiple-compressed image are remapped like those of a single compressed image, tested single, double, and triple compression, AUC = 0.52 |
| Qian and Zhang [86] | 2014 | quantisation & blocking | UCID [78](500 greyscale images, JPEG compression), QF ∈ {30, 50, 70, 80, 90} | used optimisation to analyse noise distribution, filtered grainy noise due to dithering operation, less than 22% AF images are different from uncompressed for QF = 80, 50, 30 |
| Cao *et al.* [87] | 2015 | quantisation | UCID [78] (150 uncompressed images, JPEG compression, QF ∈ {55, 65, 75, 85, 90}, overall 750 JPEG images in the experiment | Add chaotic noise dithering to quantised DCT coefficients, reduced visual quality degradation, detection accuracy = 86.31% |
| Li and Zhou [88] | 2015 | quantisation & Blocking | UCID [78] (1338 images, each compressed five times with quantisation parameter from 1 to 5) | add dither noise in prediction error domain by prediction-direction preserving strategy with minimum distortion, 100% success rate. |
| Afshin *et al.* [89] | 2016 | quantisation & blocking | UCID [78] (1338 greyscale images, partitioned into $6 \times 6$ overlapping blocks) | dictionary-based approach using a sparse representation learned by using patches extracted from the given images, efficiently reduces detection accuracy except for very high compression ratios |
| Shelke and Prasad [90] | 2018 | quantisation & blocking | UCID [78] (1300 tiff colour images, smoothing parameters ∈ {100, 150, 200, 250, 300}) | least Cuckoo search algorithm to remove detectable traces left by JPEG in both spatial and transform domains, accuracy = 97%, PSNR = 44.34 dB, MSE = 0.1789 |
| Das [91] | 2018 | blocking | BOSS database [92] (1000 random images in greyscale, 5 classical images, QF ∈ {50, 60, 70, 80, 90}) | approximation of DCT coefficients, capable of removing any distortion introduced in JPEG, unlike other AFTs requiring separate method for each, success rate = 100% with QF $\geq 60$ |
| Luo *et al.* [93] | 2018 | data-driven | BossBase V1.01 [92], BOWS2-Original [94] (10,000 uncompressed $512 \times 512$ greyscale images from each, training: 16,000, testing: 4,000 images, JPEG QF = 25, 50, 75 | GAN on TensorFlow framework running on GPU NVIDIA, superior performance without degradation in image visual quality |
| Li *et al.* [95] | 2015 | quantisation | BOSS [92] (10,000 uncompressed $512 \times 512$ images) | AFT for double JPEG compression with same quantisation matrix by slightly modifying the DCT coefficients |
| Singh and Singh [96] | 2017 | blocking | UCID [78] (1338 single and double compressed images, QF ranges from 50 to 95) | remove blocking artefacts in spatial and DCT domains, reduce grainy noise left by perceptual histogram smoothing in DCT, improved TV-based deblocking is performed in spatial domain, better tradeoff between visual quality and undetectability, high computational cost, minimum decision error against existing forensic detectors |

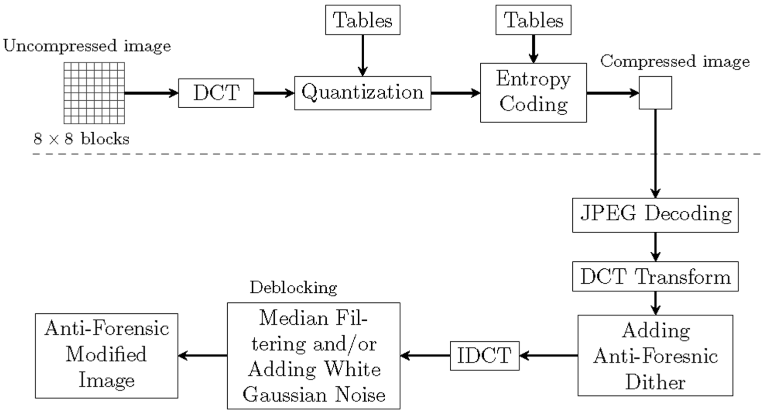

**Fig. 7** *Typical JPEG compression and AFT JPEG model*

anti-forensic dither was combined with the transform coefficients of the compressed image. The technique resulted in degradation of image quality.

The work proposed in [98] is an extension of JPEG compression anti-forensic technique presented in [77, 80] using a bitmap image as the test image. Noise is added to the DCT coefficients based on the assumption that pixel differences within and across blocks are almost the same if no compression is done. To improve the image quality, Valenzise *et al.* [99] analysed the distortions and quality degradation introduced by dithering operation in [80], and proposed an adaptive dither addition by solving a minimum-cost bipartite graph matching problem. An optimisation was proposed in [86] by analysing the noise distribution generated by [77], to filter out the grainy noise generated due to the dithering operation. Another JPEG anti-forensic scheme was proposed by Fan *et al.* [45] which adds some acceptable level of distortion to remove quantisation artefacts due to JPEG compression; yet this approach has improved image visual quality compared to [77, 80]. Similar work was also done by the same authors in [100] to fool the JPEG-based forensics while maintaining a high visual quality of the processed image. The approach was based on the non-parametric method to DCT histogram smoothing without any histogram statistical model. The results showed the effectiveness of the scheme in achieving high image visual quality while being undetectable under existing detectors. A chaos-based dither method is also presented in [87] to inject the preferred chaotic noise into the quantised DCT coefficients. The analysis showed elimination of the DCT quantisation fingerprints at the cost of small visual misrepresentations.

*2.8.2 AFTs for JPEG blocking artefacts removal:* In these techniques, the aim is to remove the blocking artefacts introduced during JPEG image compression to mislead the compression-based forensic methods. Stamm *et al.* [97] proposed AFT to suppress the JPEG blocking artefacts. The DCT coefficients were undergone smoothing operation then the addition of low power white Gaussian noise in a manner such that the Laplacian distribution of DCT coefficients is preserved. The technique was based on DCT coefficients histogram and blocking artefact measure to detect the presence of compression operation. To fool the forensic examination, it was proposed to add anti-forensic dither to the DCT coefficients after initial JPEG compression but before the second pass of JPEG compression. This will reflect as if the image has been obtained directly from a digital camera and has undergone JPEG compression within the camera just once.

Jiang *et al.* [82] eliminated blocking artefacts with low computational complexity by predicting the added noise with average detection accuracy over 99% tested on the Uncompressed Colour Image Database (UCID) [78]. Fan *et al.* [45] provided excellent AFT JPEG deblocking technique with improved visual quality and forensic undetectability compared to [80]. Fan *et al.* [83, 101] proposed a JPEG AFT to remove the traces of JPEG compression in both spatial and transform domain. A two-round TV-based deblocking and histogram smoothing operation were used to achieve a better tradeoff between image quality and hiding compression fingerprints. Qian and Zhang analysed the grainy noise distribution and proposed an improved AFT in [86] for JPEG compression using a de-noising algorithm to remove the noise due to the dithering operation and a deblocking algorithm to remove blocking artefacts. The results showed good performance when compared with [84]. Other AFT to remove the blocking artefacts in JPEG compressed images, have been proposed with improved image quality such as sparsity-based [102], knowledge evidence-based [103], fuzzy-based [104], among others.

To remove both quantisation and blocking artefacts resulting from JPEG compression, a dictionary-based method was proposed in [89]. The dictionary used a sparse representation and was learned using patches extracted from the given images and K-SVD algorithm. The proposed method was evaluated and the results showed that it can fool some of the state-of-the-art image forensics tools. Shelke and Prasad [90] proposed a Least Cuckoo search algorithm to remove detectable traces left by JPEG compression in both the spatial and transform domains. The authors devised a new fitness function, called histogram deviation, and used it for optimisation. The experimental results performed on the UCID dataset revealed good performance in fooling forensic detectors that are based on JPEG fingerprints.

Das [91] proposed a JPEG-based AFT based on approximation of DCT coefficients. The objectives were to remove the blocking artefacts and rate of zero coefficients by approximation of DCT coefficients. The technique was capable to remove any distortion introduced in JPEG, unlike other existing AFTs which require a separate technique for each. In the proposed method, first modelling of AC coefficients of DCT is performed then the model was used to restore the values of the DCT coefficients. The experiments showed the superior performance of the proposed method when evaluated against existing forensic detectors [84, 105].

*2.8.3 AFTs for wavelet-based compression:* During the JPEG compression, DCT is computed by dividing the image into blocks, which results in blocking artefacts in the compressed image. To avoid these blocking artefacts, the DCT-based compression is replaced with wavelet-based compression (JPEG2000) where the wavelet transform is computed for the whole image for better compression and image quality. Currently, few AFTs exist to hide the tampering fingerprints arising from the wavelet-based compression. In this regard, Stamm and Liu [79] for the first time remove the tampering fingerprints by adding a dither noise to the wavelet coefficients of a compressed image. The dither was added to highly textured blocks of the compressed image so as to match the coefficient distribution of the uncompressed image. Stamm and Liu [80] proposed an AFT to smooth the distributions of DCT and DWT coefficients by adding dither noise in the transform domain. The results achieved good performance on [84].

*2.8.4 AFTs for double compression:* The image after modifications or tampering is saved again in a compressed format. Detection of double JPEG compression is also widely used in image forensics. Huang *et al.* [106] showed that the values of DCT coefficients decrease with multiple compressions using the same quantisation matrix. To predict the double JPEG compression in an image, a random permutation-based algorithm was proposed. Counterfeiting the JPEG images also require re-compression, i.e. saving the tampered image as JPEG with different compression quality factor after tampering, which might provide evidence of double JPEG compression. In the literature, various double JPEG compression AFTs were proposed. The work proposed in [80] can be used to make double JPEG image compression untraceable to forensic methods based on compression history. However, due to the addition of dithering noise, the image quality is degraded.

Similarly, Sutthiwan and Shi [81] proposed an AFT to hide the traces of double JPEG compression with different quantisation matrices via 'Shrink-and-Zoom'. Feng *et al.* [107] and Milani *et al.* [108] both proposed to eliminate the double quantisation fingerprints with a good disguise effect on the tampered images by using the same concept based on Benford's law of first digit distribution by altering the first digit probability mass function to match that of a single compressed image. Qian and Qiao [109] applied enhancement and filtering operations to hide the double JPEG compression traces of comb-like DCT and blocking artefacts with better image quality. The technique easily fools the forgery detection techniques based on quantisation estimation and global histogram analysis.

Similarly, Pedro and Fernando [110] revisited the problem of attacking histogram-based forensic tools [36] and for Benford's law-based detectors, the modification distortion optimisation problem was reframed as a linear programming problem. Therefore, well-known optimisation tools like simplex were used to find the optimal solution. The technique was of heuristic nature and sub-optimal. Pasquini *et al.* [111] proposed AFT for JPEG forensics through modifying the distribution of the first significant digit (FSD) of the DCT coefficients. The FSD distribution of either a single/double compressed image was restored to that of an uncompressed image or a double compressed image to that of a single compressed image in order to fool the JPEG forensic detectors.

*2.8.5 Generative adversarial network based:* Recently, Luo *et al.* [93] adopted the framework of a GAN [112] to counter the JPEG-based image forensics. The GAN framework comprised of a generator and a discriminator, where the generator learnt itself to conceal the JPEG traces during the optimisation process against the discriminator. The results showed superior performance in a data-driven way against state-of-the-art JPEG-based AFTs without degradation in visual quality. A comprehensive analysis of adversarial attacks on CNN-based image forensic detection algorithms is provided by Gragnaniello *et al.* [113]. The analysis consists of testing the performance in terms of hiding various traces left by various manipulations during the CNN training.

## 3 Anti-anti-forensics techniques

Image forensics has gained a lot of research interest during the last few years, due to its importance in image authentication. In the tampering process, there are some inconsistencies left in the image. The traces left in the image during forgery operation will be exploited in tampering detection. However, knowing the type of traces used in tampering detection, the forger may exploit the forensics methods to hide these traces in a more convincing way known as anti-forensics. Moreover, to rebuild more credibility in forensics, several approaches have been proposed to counter anti-forensics and known as anti-anti-forensics. In this regard, many counter measures against anti-forensics have also been proposed [114]. Based on the employed methodology, counter image anti-forensic techniques are grouped into different categories and discussed in details in the following subsections.

### 3.1 AAFTs for median filtering

Zeng *et al.* [115] proposed a method to counter the anti-forensics of median filtering [44] based on the features calculated from the adjacent pixel differences in horizontal directions for each row and ratios of these differences with zero values. The periodicity of Fourier transform of these ratios for each row exposed image anti-forensic operations. The experiments were conducted on two benchmark image forensics databases UCID [78] and Dresdon Image Database [116]. Similarly, Kang *et al.* [117] discovered that the addition of noise in image pixels in AFTs for median filtering, leave visible clues in the histogram of Median Filtering Residual (MFR). These traces can be captured using ratios of histogram bins of the MFR in the textured regions. The traces left in the MFR were eliminated by the addition of noise to image pixels while maintaining least SNR.

### 3.2 AAFTs for LCA

The LCA is an imaging trace present in optical imaging systems produced due to the lens inabilities in focusing all wavelengths from a single point source in a scene to a single focal point on the sensor. It creates imperceptible colour fringes around object edges in an image. The localised inconsistencies in LCA are strong fingerprints for image splicing (i.e. cut-and-paste forgery) detection [55]. Mayer and Stamm [118], for the first time, proposed an AAF technique to counter the traces left by the LCA anti-forensics (i.e. resizing differences between colour channels) [56]. The features were extracted based on spectral properties of resampling artefacts, amplitude and phase angle differences across the colour channels at frequencies related to JPEG blocking discontinuities. Besides this, localisation of LCA-based anti-forensically manipulated image patches was carried out using statistical tests. The results showed that the algorithm performance depends on the patch size and anti-forensic scaling factor.

### 3.3 AAFTs for CFA

Chen *et al.* [119] proposed an AAFT to detect the anti-forensic operation used to falsify source camera information in an image. The method characterised various content-independent local pixel relationships that may be introduced by either authentic demosaicing methods or anti-forensic attacks. Using features from diverse demosaicing residuals and CFA-aware geometric pattern, the authors carried out some experiments and demonstrated 99% detection accuracy in revealing the anti-forensic attacks, robustness to JPEG compression, and effectively identify the true source camera model from an anti-forensic modified image.

### 3.4 AAFTs for resampling

Regarding the counter anti-forensics of resampling, Cao *et al.* [120] introduced a semi-nonintrusive approach to detect anti-forensics of resampling [71] for specially designed images. Similarly, Peng *et al.* [121] showed that irregular sampling followed by interpolation in anti-forensics methods changes the linear relationships among the neighbouring pixels. To counter this, features were derived from the auto-correlation coefficients of neighbouring pixels. A support vector machine (SVM) classifier using a Gaussian kernel was used for classification of authentic and resampled images. The results were validated on VelD database [78] containing 1338 grey images. The optimal parameters of SVM were calculated using a grid-search approach and five-fold cross-validation on the training set.

The same authors in [122] proposed an effective and secure detector to simultaneously detect resampling and its forged resampling due to some anti-forensic operations. It was observed that interpolation operations used in the resampling and forged resampling showed different statistical behaviours from the unaltered images, especially in the high-frequency domain. To capture interpolation traces, multi-directional high-pass filters were applied to both image itself and its residual to create multi-directional differences. Then, these differences were fit into an AR model. Finally, the AR coefficients and normalised histograms of the differences were used as the features. These features fed into an SVM classifier to detect both resampling and forged resampling. Experiments conducted on a large image database showed the effectiveness of the method. Furthermore, compared to the state-of-the-art, the detector achieved significant improvements in the detection of down-sampling or resampling under JPEG compression.

### 3.5 AAFTs for compression

Most of the anti-forensic techniques rely on dithering operations [77, 80, 97], where noise dithering is mostly used to hide the JPEG compression traces. Table 3

shows a summary of several methods. Lai and Böhme [105] proposed two methods to expose anti-forensics in JPEG images that modifies DCT coefficients to smooth the marginal distributions of the AC subbands. The first method works on the assumption that an anti-forensic operation does not alter high-frequency DCT coefficients. On the contrast, the second detection method exploits the difference between the variance in DCT coefficients of a given image and its corresponding tampered one. The authors demonstrated the effectiveness of both techniques against the anti-forensic attacks described in [77]. Other methods to counter the JPEG compression anti-forensic techniques were proposed, e.g. in [130] based on the noise level estimation, and in [123, 133] based on machine learning.

Bhardwaj and Pankajakshan [131] proposed AAFT for JPEG images by exploiting the inter-block correlation in DCT coefficients. The results showed this technique is effective for images manipulated with aligned and non-aligned JPEG AFTs. Bhardwaj *et al.* in [132] exploited the concept that JPEG anti-forensic operation changes the values of DCT coefficients and more decrease is observed especially in high-frequency subbands. A normalised feature was calculated as a normalised difference between absolute values of DCT coefficients in 28 high-frequency AC-subbands of the test image and its anti-forensically modified version. Assume the original image in JPEG format compressed with a quality factor $Q_r$ and its modified version are represented by $X$ and $X_m$, respectively, and their corresponding block-DCT transforms are represented by $X_d$ and $X_{md}$, respectively. Then, the sum of absolute values of DCT coefficients for each of the 28 high-frequency subbands (in zig-zag scan order) is calculated for both $X_d$ and $X_{md}$, and is denoted by $S_d^i$ and $S_{md}^i$, respectively. The normalised difference ($D$) is calculated by

$$D = \sum_i \left| \frac{S_d^i - S_{md}^i}{S_{md}^i} \right| \quad (2)$$

where $i$ represents subband index for $i = 37, 38, \ldots, 64$. The power of the proposed AAFT in differentiating between uncompressed and anti-forensically modified JPEG images is demonstrated empirically.

Existing anti-forensic methods are able to fool the JPEG-based image forensic methods based on the analysis of first-order statistics or image histogram. Singh and Singh [135] proposed to use second-order statistical features derived from the co-occurrence matrices to detect the JPEG anti-forensic footprints. The features are then used to train a SVM classifier. The results on benchmark databases showed that this scheme is capable of detecting JPEG

**Table 3** Summary of JPEG counter AFTs in chronological order

| Reference | Year | Target | Experimental Setting | Remarks |
|---|---|---|---|---|
| Lai and Böhme [105] | 2011 | quantisation | UCID [78], 800 images, QF = 60, 70, 80, 90 | pretty good results to detect JPEG compression anti-forensics |
| Li *et al.* [123] | 2012 | quantisation & Blocking | UCID [78] (1338 images, JPEG QF ∈ [49, 90], 100 Features, SVM) | based on machine learning, detection accuracy above 99% |
| Valenzise *et al.* [124] | 2013 | quantisation | UCID [78] (1338 greyscale 512 × 384 images, luma channels), NRCS [125] (978 images of size 2100 × 1500 downsampled to same size as UCID) | measures grainy noise due to anti-forensic dither in DCT domain using TV, accuracy almost 100% over QF ∈ [30, 84] |
| Wang *et al.* [126] | 2014 | quantisation | UCID [78](1338 greyscale 512×384 images, compression rate from 0.5 to 8bpp, SVM, JPEG2000, SPIHT) | analyse relationships of DWT coefficients across different decomposition levels, apply Hough transform to joint DWT histogram to derive features for SVM, AUC = 0.97 |
| Fahmi and Würtz [127] | 2016 | quantisation | own dataset of 280 images, QF ∈ [30, 90] | noise dithering is measured by spatial frequency phase variations in tampered image blocks, highest accuracy = 93 at QF = 30 and 70% for QF = 30 and 95 |
| Barni *et al.* [128] | 2016 | double JPEG compression | RAISE dataset [129] (1400 images for training, 300 for kernel settings, 300 for testing, size 4288 ×2848 and 2144×1424, QF ∈ [49, 87]) | adversary-aware AAFT for double JPEG compression based on data-driven approaches and adversary-aware SVM, efficiently counter universal anti-forensic attacks |
| Zeng *et al.* [130] | 2018 | quantisation | 1338 greyscale images from UCID [78] and 10,000 images from BOSSbase [92] compressed at QF = 65, 75, 85, 95 | based on noise level estimation, thresholding functions, simple to implement, excellent performance on various databases, detection rate = 92.86% |
| Bhardwaj and Pankajakshan [131] | 2018 | blocking | own dataset with 4338 greyscale images (3000 from BossBase v1.01 and 1338 from UCID compressed at QF = 40, 50, 60, 70, 80 | based on inter-block correlations in DCT subbands, 99% images detected as anti-forensically modified |
| Bhardwaj *et al.* [132] | 2018 | quantisation | 3691 greyscale images (1338 from UCID [78], 353 from Dresden [116] and 2000 from BossBase v1.01 [92]) compressed with QF = 40, 50, 60, 70, 80 | normalised difference features are calculated for measuring the changes in DCT coefficients by anti-forensic operation, AUC > 90% for three databases and QF = [39, 49, 59, 69, 80] |
| Li *et al.* [133] | 2018 | quantisation & blocking | 10,000 greyscale images from BossBase v1.0 [92] cropped from centre, downsampled to $512 \times 512$ | based on machine learning, computational expensive, detection performance drops after JPEG compression |
| Li *et al.* [134] | 2019 | double JPEG compression | convolution kernels initialised with zero-mean Gaussian distribution with 0.01 standard deviation, bias parameters initialised with uniform distribution ∈ [0, 1], FCN layers, 'Xavier' initialisation, bias parameters set to zeros | CNN based, promising detection performance without designing specific features for each anti-forensic operation, accuracy > 99% in classifying single and double JPEG compression at different QFs |

compression for the case when the anti-forensic operation is applied to hide the traces of tampering. The algorithm is also able to counter CE and median filtering anti-forensics, as well as detecting various image processing operations like scaling, rotation, filtering (mean filtering, Gaussian filtering, Weiner filtering) and proved to be used as a multipurpose anti-forensic countering method.

*3.5.1 Phase-based JPEG AAFT:* Fahmi *et al.* [127, 136] proposed an AAFT to detect the presence of noise dithering using the spatial frequency, phase variations in tampered image blocks.

*3.5.2 Using image quality features:* Valenzise *et al.* [124] proposed a method to counter anti-forensics of JPEG compression [80] using image quality features and TV. The core of the detector was based on the 'Recompress and Observe' paradigm. It was an extension of [137] to the most general and challenging scenario where the quantisation matrix is unknown. It was demonstrated that adding noise dither in DCT coefficients to hide tampering traces introduces grainy noise in the spatial domain. The authors estimated the noise in the recompressed images by means of TV, i.e. the $\ell_1$ norm of the spatial first-order derivatives [138].

*3.5.3 Data fusion:* Fontani *et al.* [139] detected the anti-forensic operation on JPEG compressed images with an assumption of AFTs applied to the image. A binary classifier is used to differentiate between images with anti-forensics methods and untampered images.

*3.5.4 Noise estimation based:* To counter the anti-forensic of JPEG compression (i.e. removal of quantisation and blocking artefacts), an AAFT based on the noise level estimation was proposed by [130]. The technique was implemented using thresholding functions and simple to implement. The results showed excellent performance in various databases.

*3.5.5 AAFTs for wavelet-based compression:* Wang *et al.* [126] proposed an AAFT for wavelet-based compression. The fingerprints of counter-forensics wavelet-based compression were suppressed by analysing relationships of DWT coefficients across different decomposition levels. Hough transform was applied to the joint DWT histogram to derive the feature vector using a SVM.

*3.5.6 Double JPEG compressed based:* In the literature, there also exist anti-forensic techniques to hide the double compression fingerprints. In this regard, Barni *et al.* [128] proposed an adversary-aware AAFT for double JPEG compression based on data-driven approaches and adversary-aware SVM. The results showed effectiveness in countering the universal anti-forensic attacks. Similarly, Li *et al.* [134] also described another data-driven approach to detect double JPEG compression and its related anti-forensic operations by using a CNN. The CNN inputs are raw JPEG DCT coefficients and decompressed image pixels. The technique showed promising performance in addressing the counter anti-forensic issues without designing specific features for each anti-forensic operation.

### *3.6 AAFTs for CE*

*3.6.1 Statistics based:* The CE detectors based on first-order statistics are shown to be less robust. Lin *et al.* [30] used the inter-channel dependency due to colour image interpolation to detect CE in an image. De Rosa *et al.* [37] proposed a technique to counter the anti-forensic of CE. The authors used the GLCM of an image to determine whether it is CE manipulated or not. The use of second-order statistics of the image is shown to be more effective against anti-forensic of CE targeted at histogram-based detectors. Stamm and Liu [140] proposed an iterative algorithm to jointly estimate the histogram of the original image as well as the mapping function used for CE from the contrast-enhanced image. The results showed the effectiveness of the algorithm in accurately estimating the CE mapping function and histogram of the unaltered image.

*3.6.2 CNN based:* Sun *et al.* [141] and Yang *et al.* [142] proposed the use of a CNN for detecting CE to counter the anti-CE and JPEG compression attacks. The authors demonstrated with experiments that CE fingerprints are destroyed when an image is JPEG compressed after CE. However, the residual traces remain in GLCM image and are used as an input to the CNN layer. The results showed that this technique can efficiently detect CE even when the image is JPEG compressed. Similarly, Shan *et al.* [143] proposed a robust JPEG forensic technique based on CNN to detect both local and global CE. Experimental results showed the technique efficiently detect both local and global CE in compressed images irrespective of the order of CE and JPEG compression.

### *3.7 AAFTs for HE*

During the process of image tampering, some filtering and enhancement operations (such as median filtering, gamma correction, HE, and noise reduction) are applied to conceal the traces of tampering and to maintain a more realistic look of tampered images. Existing image forensic methods fail to detect HE of low-resolution and highly compressed images. Akhtar and Khan [144] presented a technique to determine whether a HE operation is applied to a given image (uncompressed or compressed) or not. The authors analysed the image histogram in the frequency domain and exploited the difference in DC and AC coefficients in histogram's discrete Fourier transform. This technique outperformed existing methods that detect HE operation even if the image is JPEG compressed after the equalisation operation. However, it works only for a standard HE operation and does not work for other advanced equalisation techniques like contrast limited adaptive HE, dynamic HE etc.

### *3.8 Multipurpose AAF techniques*

The aforementioned AAFTs aim to detect or counter specifically one class of anti-forensics and are not extendable for other types of anti-forensics. Recently, some research has been conducted to provide a multipurpose framework for countering various types of anti-forensics without prior knowledge of the anti-forensics operations which have been utilised. These approaches are discussed as follows.

*3.8.1 Using AR model:* Zeng *et al.* [145] proposed a multipurpose countermeasure by exploiting the trails neglected by five widely known anti-forensic techniques. This analysis revealed that the application of these AFTs disturbs the neighbouring pixels relationship in the original images. Hence, the proposed countermeasure used an AR model which can characterise the correlations between neighbouring pixels with fewer features (AR coefficients) than other techniques such as Subtractive Pixel Adjacency Matrix [146] and Local Binary Pattern [39]. The authors demonstrated comparable results, or even better, to those attained using AF-specific countermeasures.

*3.8.2 Deep learning:* Deep learning approaches have been found successful in vast image processing applications. Unlike traditional approaches which are based on hand-crafted features which are laborious and require domain knowledge, deep learning can learn to extract important features automatically from raw images. Yu *et al.* [147] proposed a framework using a CNN to counter multiple anti-forensics. Based on the automatic feature extraction, the approach was capable of detecting various kind of anti-forensics attacks in a tampered image and showed superior performance compared to well-known image AFTs. Chen *et al.* [148] proposed another effective CNN architecture to detect manipulated images under anti-forensic attacks. In [149], the authors proposed a generalised 3D CNN for face anti-spoofing taking into consideration both spatial and temporal information.

## 4 Discussions and conclusions

In this paper, we presented a critical review of anti-forensics and AAFTs. The purpose is to facilitate for researchers interested in the field of digital image forensics to get an insight into state-of-the-art AFTs and AAFTs. People from both academia and industry can also benefit from this survey study to identify research gaps in order to develop more robust digital image forgery detection techniques to counter the attacks of anti-forensics. Countering AFTs in digital images is an extremely important and challenging endeavour in our modern society and limited work has been done in this area and can be considered a prominent research direction for people interested image forensics.

Through this review, it is emphasised that the battle never ends between forensic, anti-forensic and counter anti-forensic techniques. New computational algorithms and models are always developed to hide the traces of tampering day-by-day which need to be discovered by others. Despite the remarkable work that has been done in digital image forensic in general, the image acquisition and manipulation technologies continue to evolve and tomorrow's counter forensic techniques will definitely differ, affecting people's trust in visual content. There is a lot of research potential to handle diverse situations and develop more robust and general-purpose computational techniques to identify fake images that undergo subtle alteration. For instance, it is demonstrated that adding dither to image pixels have shown effective results in hiding traces of images compression, CE, and unsharp masking. It would be great if a single framework is developed to handle the anti-forensics for different tampering operations based on dithering noise addition. Anti-forensics, however, has shaken the confidence in multimedia forensics. It is therefore extremely important to secure images against anti-forensics attacks by detecting fingerprints of left traces by the AFTs themselves. In this regard, research has been started to exploit various traces to counter the anti-forensics attacks. Similar to the targeted forensic methods that aim at detecting a certain operation, though the countering methods of anti-forensic can effectively expose the resulting images after some anti-forensic operations, their applicability is still limited since each of them just focuses on a specific anti-forensic operation.

The study on image anti-forensics highlighted the importance of the trustworthiness of various image forensic methods being used. From the study, we conclude that image forensic attacks are dependent on image forensic techniques. The image forensics based on a single method is affected by the anti-forensic method. Therefore, it is important to use different forensic methodologies to capture the various forensic fingerprints. Since, it is difficult to conceal the different tampering traces simultaneously. The anti-forensic operations also leave some new traces that can effectively counter the anti-forensic operations. A lot of research potential is foreseen in image forensics field to prove the trustworthiness of digital images. A suggested future research direction is to investigate the role of image quality while hiding the traces of image forensics and anti-forensics in AFTs and AAFTs, respectively. Image forensics is a subset of the forensic discipline in computer science. Later, several fields are evolved like video forensics, disk forensics, network forensics, computer forensics, printer forensics etc. It would be interesting to extend the work on digital images anti-forensics and counter anti-forensics for the before mentioned evolving fields of digital forensics.

Our analysis raises a serious question about the security and robustness of these forensics algorithms because a single anti-forensics strategy is able to defeat at least three different forensics methods. We also observe that the distortion measures used by these forensics algorithms are somehow related, as minimisation of one also effectively reduces the others. Though, the extent of reduction varies from one measure to another. In fact, if we can establish the relationships among different image artefact measures, more secure forensics technique can be designed.

There are several other potential research directions in this area. For instance, it is required to develop a comprehensive benchmark system that can support decision makers to select among various anti-anti-forensic techniques under diverse scenarios. Despite existing individual efforts, one of the obstacles to evaluate the performance of a variety of methods is the lack of reliable large datasets that cover different anti-forensic techniques. Moreover, datasets are naturally imbalanced with large number of images being original and few being forged. Machine learning algorithms dealing balanced datasets will perform poorly in classifying these imbalanced datasets. Two more key research areas that have great potential in the future are the rapid advancement of multi-sensor mega-pixel cameras and deep learning technologies. This requires re-evaluating existing forensic techniques and designing more innovative methods to fuse multiple models and handle massive amount of data on limited-resources devices. The rise of cloud computing systems can provide a scalable powerful platform to leverage digital image forensics and anti-anti-forensics technologies.

## 5 References

[1] Stamm, M.C., Wu, M., Liu, K.J.R.: 'Information forensics: an overview of the first decade', *IEEE. Access.*, 2013, **1**, pp. 167–200

[2] Zhu, B.B., Swanson, M.D., Tewfik, A.H.: 'When seeing isn't believing [multimedia authentication technologies]', *IEEE Signal Process. Mag.*, 2004, **21**, (2), pp. 40–49

[3] Piva, A.: 'An overview on image forensics', *ISRN Signal Process.*, 2013, **2013**, pp. 1–22

[4] Zheng, L., Zhang, Y., Thing, V.L.: 'A survey on image tampering and its detection in real-world photos', *J. Vis. Commun. Image Represent.*, 2019, **58**, pp. 380–399

[5] Korus, P.: 'Digital image integrity – a survey of protection and verification techniques', *Digit. Signal Process.*, 2017, **71**, pp. 1–26

[6] Warif, N.B.A., Wahab, A.W.A., Idris, M.Y.I., *et al.*: 'Copy-move forgery detection: survey, challenges and future directions', *J. Netw. Comput. Appl.*, 2016, **75**, pp. 259–278

[7] Johnston, P., Elyan, E.: 'A review of digital video tampering: from simple editing to full synthesis', *Digit. Invest.*, 2019, **29**, pp. 67–81

[8] Rocha, A., Scheirer, W., Boult, T., *et al.*: 'Vision of the unseen: current trends and challenges in digital image and video forensics', *ACM Comput. Surv. (CSUR)*, 2011, **43**, (4), pp. 26:1–26:42

[9] Birajdar, G.K., Mankar, V.H.: 'Digital image forgery detection using passive techniques: a survey', *Digit. Invest.*, 2013, **10**, (3), pp. 226–245

[10] Qureshi, M.A., Deriche, M.: 'A bibliography of pixel-based blind image forgery detection techniques', *Signal Process., Image Commun.*, 2015, **39**, pp. 46–74

[11] Dixit, R., Naskar, R.: 'Review, analysis and parameterisation of techniques for copy–move forgery detection in digital images', *IET Image Process.*, 2017, **11**, (9), pp. 746–759

[12] Lin, X., Li, J.H., Wang, S.L., *et al.*: 'Recent advances in passive digital image security forensics: a brief review', *Engineering*, 2018, **4**, (1), pp. 29–39

[13] Qureshi, M.A., Deriche, M.: 'A review on copy move image forgery detection techniques'. 11th Int. Multi-Conf. on Systems, Signals & Devices (SSD). IEEE, Barcelona, Spain, 2014, pp. 1–5

[14] Guo, Y., Cao, X., Zhang, W., *et al.*: 'Fake colorized image detection', *IEEE Trans. Inf. Forensics Sec.*, 2018, **13**, (8), pp. 1932–1944

[15] Mahmood, T., Nawaz, T., Mehmood, Z., *et al.*: 'Forensic analysis of copy-move forgery in digital images using the stationary wavelets'. Proc. 6th IEEE Int. Conf. on Innovative Computing Technology (INTECH), Dublin, Ireland, 2016, pp. 578–583

[16] Farid, H.: 'Image forgery detection: a survey', *IEEE Signal Process. Mag.*, 2009, **26**, (2), pp. 16–25

[17] Ansari, M.D., Ghrera, S.P., Tyagi, V.: 'Pixel-based image forgery detection: a review', *IETE J. Educ.*, 2014, **55**, (1), pp. 40–46

[18] Redi, J.A., Taktak, W., Dugelay, J.L.: 'Digital image forensics: a booklet for beginners', *Multimedia Tools Appl.*, 2011, **51**, (1), pp. 133–162

[19] Dong, J., Wang, W., Tan, T.: 'CASIA image tampering detection evaluation database'. IEEE China Summit & Int. Conf. on Signal and Information Processing (ChinaSIP). IEEE, 2013, pp. 422–426

[20] Zhu, X., Qian, Y., Zhao, X., *et al.*: 'A deep learning approach to patch-based image inpainting forensics', *Signal Process., Image Commun.*, 2018, **67**, pp. 90–99

[21] Qureshi, M.A., Deriche, M., Beghdadi, A., *et al.*: 'A critical survey of state-of-the-art image inpainting quality assessment metrics', *J. Vis. Commun. Image Represent.*, 2017, **49**, pp. 177–191

[22] Pranita, D., Monika, R.: 'Survey on anti-forensics operations in image forensics', *Int. J. Comput. Sci. Inf. Technol.*, 2014, **5**, pp. 1570–1573

[23] Singh, N., Joshi, S.: 'Digital image forensics and counter antiforensic'. Proc. of Int. Conf. on Recent Cognizance in Wireless Communication and Image Processing, Springer, 2016, pp. 805–811

[24] Gül, M., Kugu, E.: 'A survey on anti-forensics techniques'. IEEE Int. Artificial Intelligence and Data Processing Symp. (IDAP), Malatya, Turkey, 2017, pp. 1–6

[25] Shelke, P.M., Prasad, R.S.: 'Tradeoffs between forensics and anti-forensics of digital images'. Computer Vision: Concepts, Methodologies, Tools, and Applications. IGI Global, IGI Global, Pennsylvania, USA, 2018, pp. 2124–2138

[26] Böhme, R., Kirchner, M.: 'Counter-forensics: attacking image forensics'. Digital Image Forensics, New York, 2013, pp. 327–366

[27] Cao, G., Zhao, Y., Ni, R., *et al.*: 'Contrast enhancement-based forensics in digital images', *IEEE Trans. Inf. Forensics Sec.*, 2014, **9**, (3), pp. 515–525

[28] Stamm, M.C., Liu, K.J.R.: 'Blind forensics of contrast enhancement in digital images'. IEEE Int. Conf. on Image Processing, San Diego, CA, USA, 2008, pp. 3112–3115

[29] Stamm, M.C., Liu, K.J.R.: 'Forensic detection of image manipulation using statistical intrinsic fingerprints', *IEEE Trans. Inf. Forensics Secur.*, 2010, **5**, (3), pp. 492–506

[30] Lin, X., Li, C.T., Hu, Y.: 'Exposing image forgery through the detection of contrast enhancement'. Proc. 20th IEEE Int. Conf. on Image Processing (ICIP), Melbourne, VIC, Australia, 2013, pp. 4467–4471

[31] Cao, G., Zhao, Y., Ni, R., *et al.*: 'Anti-forensics of contrast enhancement in digital images'. Proc. of the 12th ACM Workshop on Multimedia and Security (MM&Sec), Rome, Italy, 2010, pp. 25–34

[32] Cao, G., Zhao, Y., Ni, R., *et al.*: 'Attacking contrast enhancement forensics in digital images', *Sci. China Inform. Sci.*, 2014, **57**, (5), pp. 1–13

[33] Kwok, C.W., Au, O.C., Chui, S.H.: 'Alternative anti-forensics method for contrast enhancement'. 10th Int. Workshop on Digital Forensics and Watermarking, (vol. 7128 LNCS, Atlantic City, NY, USA, 2011), pp. 398–410

[34] Barni, M., Fontani, M., Tondi, B.: 'A universal technique to hide traces of histogram-based image manipulations'. Proc. 14th ACM Workshop on Multimedia and Security (MM&Sec), Coventry, UK, 2012, pp. 97–104

[35] Barni, M., Fontani, M., Tondi, B.: 'A universal attack against histogram-based image forensics', *Int. J Dig Crime Forens*, 2013, **5**, (3), pp. 35–52

[36] Comesana.Alfaro, P., Pérez.González, F.: 'Optimal counterforensics for histogram-based forensics'. Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 2013, pp. 3048–3052

[37] De Rosa, A., Fontani, M., Massai, M., *et al.*: 'Second-order statistics analysis to cope with contrast enhancement counter-forensics', *IEEE Signal Process. Lett.*, 2015, **22**, (8), pp. 1132–1136

[38] Ravi, H., Subramanyam, A., Emmanuel, S.: 'ACE – an effective anti-forensic contrast enhancement technique', *IEEE Signal Process. Lett.*, 2016, **23**, (2), pp. 212–216

[39] Ding, F., Zhu, G., Shi, Y.Q.: 'A novel method for detecting image sharpening based on local binary pattern'. Int. Workshop on Digital Watermarking, Auckland, New Zealand, 2013, pp. 180–191

[40] Cao, G., Zhao, Y., Ni, R., *et al.*: 'Unsharp masking sharpening detection via overshoot artifacts analysis', *IEEE Signal Process. Lett.*, 2011, **18**, (10), pp. 603–606

[41] Laijie, L., Gaobo, Y., Ming, X.: 'Anti-forensics for unsharp masking sharpening in digital images', *Int. J. Digit. Crime Forensics (IJDCF)*, 2013, **5**, (3), pp. 53–65

[42] Kang, X., Stamm, M.C., Peng, A., *et al.*: 'Robust median filtering forensics using an autoregressive model', *IEEE Trans. Inf. Forensics Sec.*, 2013, **8**, (9), pp. 1456–1468

[43] Nguyen, D.T., Gebru, I., Conotter, V., *et al.*: 'Counter-forensics of median filtering'. IEEE 15th Int. Workshop on Multimedia Signal Processing (MMSP), Pula, Italy, 2013, pp. 260–265

[44] Wu, Z.H., Stamm, M.C., Liu, K.R.: 'Anti-forensics of median filtering'. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, BC, 2013, pp. 3043–3047

[45] Fan, W., Wang, K., Cayre, F., *et al.*: 'Median filtered image quality enhancement and anti-forensics via variational deconvolution', *IEEE Trans. Inf. Forensics Sec.*, 2015, **10**, (5), pp. 1076–1091

[46] Fontani, M., Barni, M.: 'Hiding traces of median filtering in digital images'. Proc. of the 20th European Signal Processing Conf. (EUSIPCO), Bucharest, Romania, 2012, pp. 1239–1243

[47] Chen, C., Ni, J., Huang, J.: 'Blind detection of median filtering in digital images: a difference domain based approach', *IEEE Trans. Image Process.*, 2013, **22**, (12), pp. 4699–4710

[48] Yuan, H.D.: 'Blind forensics of median filtering in digital images', *IEEE Trans. Inf. Forensics Sec.*, 2011, **6**, (4), pp. 1335–1345

[49] Zhang, Y., Li, S., Wang, S., *et al.*: 'Revealing the traces of median filtering using high-order local ternary patterns', *IEEE Signal Process. Lett.*, 2014, **21**, (3), pp. 275–279

[50] Sharma, S., Subramanyam, A.V., Jain, M., *et al.*: 'Anti-forensic technique for median filtering using L1-L2 TV model'. 8th IEEE Int. Workshop on Information Forensics and Security (WIFS), Abu Dhabi, United Arab Emirates, 2016, pp. 1–6

[51] Kim, D., Jang, H.U., Mun, S.M., *et al.*: 'Median filtered image restoration and anti-forensics using adversarial networks', *IEEE Signal Process. Lett.*, 2018, **25**, (2), pp. 278–282

[52] Johnson, M.K., Farid, H.: 'Exposing digital forgeries through chromatic aberration'. Proceeding of the 8th Workshop on Multimedia and Security, Geneva, Switzerland, 2006, pp. 48–55

[53] Mayer, O., Stamm, M.: 'Improved forgery detection with lateral chromatic aberration'. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 2016, pp. 2024–2028

[54] Yerushalmy, I., Hel.Or, H.: 'Digital image forgery detection based on lens and sensor aberration', *Int. J. Comput. Vis.*, 2011, **92**, (1), pp. 71–91

[55] Mayer, O., Stamm, M.C.: 'Accurate and efficient image forgery detection using lateral chromatic aberration', *IEEE Trans. Inf. Forensics Sec.*, 2018, **13**, (7), pp. 1762–1777

[56] Mayer, O., Stamm, M.C.: 'Anti-forensics of chromatic aberration'. Media Watermarking, Security, and Forensics, SPIE, San Francisco, California, USA, 2015, vol. 9409, pp. 94090M–94090M

[57] Sengupta, P., Sameer, V.U., Naskar, R., *et al.*: 'Source anonymization of digital images: a counter–forensic attack on PRNU based source identification techniques'. Annual ADFSL Conf. on Digital Forensics, Security and Law, Florida, USA, 2017, pp. 95–106

[58] Villalba, L.J.G., Orozco, A.L.S., Corripio, J.R., *et al.*: 'A PRNU-based counter-forensic method to manipulate smartphone image source identification techniques', *Future Gener. Comput. Syst.*, 2017, **76**, pp. 418–427

[59] Tuama, A., Comby, F., Chaumont, M.: 'Camera model identification with the use of deep convolutional neural networks'. Proc. of IEEE Int. Workshop on Information Forensics and Security (WIFS), Abu Dhabi, United Arab Emirates, 2016, pp. 1–6

[60] Chen, C., Zhao, X., Stamm, M.C.: 'MISLGAN: an anti-forensic camera model falsification framework using a generative adversarial network'. Proc. of the 25th IEEE Int. Conf. on Image Processing (ICIP), Athens, Greece, 2018, pp. 535–539

[61] Dirik, A.E., Memon, N.: 'Image tamper detection based on demosaicing artifacts'. 16th IEEE Int. Conf. on Image Processing (ICIP), Cairo, Egypt, 2009, pp. 1497–1500

[62] Gallagher, A.C., Chen, T.: 'Image authentication by detecting traces of demosaicing'. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops, Anchorage, AK, USA, 2008

[63] Böhme, R., Kirchner, M.: 'Synthesis of color filter array pattern in digital images'. Proc. SPIE, Media Forensics and Security XI, San Jose, California, USA, 2009, vol. 7254, pp. 72540K–72540K-14

[64] Guangling, S., Zhoubiao, S., Yuejun, C.: 'Color filter array synthesis in digital image via dictionary re-demosaicing'. Int. Conf. on Multimedia Information Networking and Security (MINES), Nanjing, Jiangsu, China, 2010, pp. 898–901

[65] Chuang, W.H., Wu, M.: 'Robustness of color interpolation identification against anti-forensic operations'. Information Hiding (vol. 7692 LNCS, Berkeley, CA, USA, 2012), pp. 16–30

[66] Popescu, A.C., Farid, H.: 'Exposing digital forgeries by detecting traces of resampling', *IEEE Trans. Signal Process.*, 2005, **53**, (2), pp. 758–767

[67] Gallagher, A.C.: 'Detection of linear and cubic interpolation in JPEG compressed images'. Second Canadian Conf. on Computer and Robot Vision, Victoria, BC, Canada, 2005, pp. 65–72

[68] Mahdian, B., Saic, S.: 'Blind authentication using periodic properties of interpolation', *IEEE Trans. Inf. Forensics Sec.*, 2008, **3**, (3), pp. 529–538

[69] Kirchner, M.: 'Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue'. 10th ACM Workshop on Multimedia and Security, Oxford, UK, 2008, pp. 11–20

[70] Gloe, T., Kirchner, M., Winkler, A., *et al.*: 'Can we trust digital image forensics?'. 15th Int. Conf. on Multimedia, Newyork, USA, 2007, pp. 78–86

[71] Kirchner, M., Böhme, R.: 'Hiding traces of resampling in digital images', *IEEE Trans. Inf. Forensics Sec.*, 2008, **3**, (4), pp. 582–592

[72] Thai, T.H., Cogranne, R., Retraint, F., *et al.*: 'JPEG quantization step estimation and its applications to digital image forensics', *IEEE Trans. Inf. Forensics Sec.*, 2016, **12**, (1), pp. 123–133

[73] Wang, Q., Zhang, R.: 'Double JPEG compression forensics based on a convolutional neural network', *EURASIP J. Inf. Secur.*, 2016, **23**, pp. 1–12

[74] Luo, Y.W., Huang, J.: 'Detection of quantization artifacts and its applications to transform encoder identification', *IEEE Trans. Inf. Forensics Sec.*, 2010, **5**, (4), pp. 810–815

[75] Ebrahimi, E.G., Ibrahim, S., Alizadeh, M.: 'Paint-doctored JPEG image forensics based on blocking artifacts'. Int. Conf. and Workshop on Computing and Communication (IEMCON), Vancouver, BC, Canada, 2015, pp. 1–5

[76] Singh, N., Bansal, R.: 'Analysis of Benford's law in digital image forensics'. Int. Conf. on Signal Processing and Communication (ICSC), Noida, India, 2015, pp. 413–418

[77] Stamm, M.C., Tjoa, S.K., Lin, W.S., *et al.* 'Anti-forensics of JPEG compression'. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Dallas, TX, USA, 2010, pp. 1694–1697

[78] Schaefer, G., Stich, M.: 'UCID: an uncompressed color image database'. Storage and Retrieval Methods and Applications for Multimedia 2004. vol. 5307. Int. Society for Optics and Photonics, San Jose, California, USA, 2003, pp. 472–481

[79] Stamm, M.C., Liu, K.J.R.: 'Wavelet-based image compression anti-forensics'. 17th Int. Conf. on Image Processing (ICIP), Hong Kong, China, 2010, pp. 1737–1740

[80] Stamm, M.C., Liu, K.J.R.: 'Anti-forensics of digital image compression', *IEEE Trans. Inf. Forensics Sec.*, 2011, **6**, (3), pp. 1050–1065

[81] Sutthiwan, P., Shi, Y.Q.: 'Anti-forensics of double JPEG compression detection'. Int. Workshop on Digital Watermarking, Atlantic City, NJ, 2011, pp. 411–424

[82] Jiang, Y., Zeng, H., Kang, X., *et al.*: 'The game of countering JPEG anti-forensics based on the noise level estimation'. Signal and Information Processing Association Annual Summit and Conf. (APSIPA), Kaohsiung, Taiwan, 2013, pp. 1–9

[83] Fan, W., Wang, K., Cayre, F., *et al.*: 'A variational approach to JPEG anti-forensics'. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, BC, Canada, 2013, pp. 3058–3062

[84] Fan, Z., de Queiroz, R.L.: 'Identification of bitmap compression history: JPEG detection and quantizer estimation', *IEEE Trans. Image Process.*, 2003, **12**, (2), pp. 230–235

[85] Barni, M., Fontani, M., Tondi, B.: 'Universal counterforensics of multiple compressed jpeg images'. 13th Int. Workshop on Digital Forensics and Watermarking, 2015, vol. 9023, pp. 31–46

[86] Qian, Z., Zhang, X.: 'Improved anti-forensics of JPEG compression', *J. Syst. Softw.*, 2014, **91**, (1), pp. 100–108

[87] Cao, Y., Gao, T., Sheng, G., *et al.*: 'A new anti-forensic scheme hiding the single JPEG compression trace for digital image', *J. Forensic Sci.*, 2015, **60**, (1), pp. 197–205

[88] Li, Y., Zhou, J.: 'Anti-forensics of lossy predictive image compression', *IEEE Signal Process. Lett.*, 2015, **22**, (12), pp. 2219–2223

[89] Afshin, N., Razzazi, F., Moin, M.S.: 'A dictionary based approach to JPEG anti-forensics'. IEEE 8th Int. Conf. on Intelligent Systems, (IS), Sofia, Bulgaria, 2016, pp. 778–783

[90] Shelke, P.M., Prasad, R.S.: 'An improved anti-forensics jpeg compression using least cuckoo search algorithm', *Imaging Sci. J.*, 2018, **66**, (3), pp. 169–183

[91] Das, T.K.: ‘Anti-forensics of JPEG compression detection schemes using approximation of dct coefficients’, *Multimedia Tools Appl.*, 2018, **77**, (24), pp. 31835–31854

[92] Bas, P., Filler, T., Pevný, T.: ‘Break our steganographic system: the ins and outs of organizing BOSS’. Int. workshop on information hiding, Prague, Czech Republic, 2011, pp. 59–70

[93] Luo, Y., Zi, H., Zhang, Q.*, et al.*: ‘Anti-forensics of jpeg compression using generative adversarial networks’. Proc. of the 26th IEEE European Signal Processing Conf. (EUSIPCO), Rome, Italy, 2018, pp. 952–956

[94] BOWS.: ‘The 2nd BOWS contest (break our watermarking system) was organised within the activity of the water marking virtual laboratory (Wavila) of the European network of excellence ECRYPT between the 17th of July 2007 and 17th of April, 2009’, 2009, http://bows2.ec-lille.fr/

[95] Li, H., Luo, W., Huang, J.: ‘Anti-forensics of double JPEG compression with the same quantization matrix’, *Multimedia Tools Appl.*, 2015, **74**, (17), pp. 6729–6744

[96] Singh, G., Singh, K.: ‘Improved JPEG anti-forensics with better image visual quality and forensic undetectability’, *Forensic Sci. Int.*, 2017, **277**, pp. 133–147

[97] Stamm, M.C., Tjoa, S.K., Lin, W.S.*, et al.*: ‘Undetectable image tampering through JPEG compression anti-forensics’. 17th Int. Conf. on Image Processing (ICIP), Hong Kong, 2010, pp. 2109–2112

[98] Manimurugan, S., Athira, B.: ‘A tailored anti-forensic technique for digital image applications’, *Int. J. Comput. Appl.*, 2012, **53**, pp. 14–20

[99] Valenzise, G., Tagliasacchi, M., Tubaro, S.: ‘The cost of JPEG compression anti-forensics’. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Prague, CZECH REPUBLIC, 2011, pp. 1884–1887

[100] Fan, W., Wang, K., Cayre, F.*, et al.*: ‘JPEG anti-forensics using non-parametric DCT quantization noise estimation and natural image statistics’. Proc. of ACM Information Hiding and Multimedia Security Workshop, Montpellier, France, 2013, pp. 117–122

[101] Fan, W., Wang, K., Cayre, F.*, et al.*: ‘JPEG anti-forensics with improved tradeoff between forensic undetectability and image quality’, *IEEE Trans. Inf. Forensics Sec.*, 2014, **9**, (8), pp. 1211–1226

[102] Zhao, C., Zhang, J., Ma, S.*, et al.*: ‘Reducing image compression artifacts by structural sparse representation and quantization constraint prior’, *IEEE Trans. Circuits Syst. Video Technol.*, 2017, **27**, (10), pp. 2057–2071

[103] Tang, A.W.K.K., Craft, N.: ‘Using a knowledge-based approach to remove blocking artifacts in skin images for forensic analysis’, *IEEE Trans. Inf. Forensics Sec.*, 2011, **6**, (3), pp. 1038–1049

[104] Gambhir, D., Rajpal, N.: ‘Fuzzy edge detector based blocking artifacts removal of DCT compressed images’. Int. Conf. on Circuits, Controls and Communications (CCUBE), Bengaluru, India, 2013, pp. 1–6

[105] Lai, S., Böhme, R.: ‘Countering counter-forensics: the case of JPEG compression’. Information Hiding, Prague, Czech Republic, 2011, vol. 6958, pp. 285–298

[106] Huang, F., Huang, J., Shi, Y.Q.: ‘Detecting double JPEG compression with the same quantization matrix’, *IEEE Trans. Inf. Forensics Sec.*, 2010, **5**, (4), pp. 848–856

[107] Feng, C., Xu, Z., Zheng, X.: ‘An anti-forensic algorithm of JPEG double compression based forgery detection’. Int. Symp. on Information Science and Engineering (ISISE), Shanghai, China, 2012, pp. 159–164

[108] Milani, S., Tagliasacchi, M., Tubaro, S.: ‘Antiforensics attacks to Benford's law for the detection of double compressed images’. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, BC, Canada, 2013, pp. 3053–3057

[109] Qian, Z., Qiao, T.: ‘Simplified anti-forensics of JPEG compression’, *J. Comput. (JCP)*, 2013, **8**, (10), pp. 2483–2488

[110] Pedro, C., Fernando, P.G.: ‘The optimal attack to histogram-based forensic detectors is simple(x)’. IEEE Int. Workshop on Information Forensics and Security (WIFS), Atlanta, GA, USA, 2014, pp. 137–142

[111] Pasquini, C., Comesana.Alfaro, P., Pérez.González, F.*, et al.*: ‘Transportation-theoretic image counterforensics to first significant digit histogram forensics’. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 2014, pp. 2699–2703

[112] Barni, M., Stamm, M.C., Tondi, B.: ‘Adversarial multimedia forensics: overview and challenges ahead’. Proc. of the 26th IEEE European Signal Processing Conf. (EUSIPCO), Rome, Italy, 2018, pp. 962–966

[113] Gragnaniello, D., Marra, F., Poggi, G.*, et al.*: ‘Analysis of adversarial attacks against CNN-based image forgery detectors’. Proc. of 26th European Signal Processing Conf. (EUSIPCO), Rome, Italy, 2018, pp. 967–971

[114] Singh, N., Joshi, S.: ‘Digital image forensics and counter anti-forensics’. Proc. of the Int. Conf. on Recent Cognizance in Wireless Communication & Image Processing, 2016, pp. 805–811

[115] Zeng, H., Qin, T., Kang, X.*, et al.*: ‘Countering anti-forensics of median filtering’. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Springer, 2014, pp. 2704–2708

[116] Gloe, T., Böhme, R.: ‘The ‘Dresden image database’ for benchmarking digital image forensics’. Proc. of the 2010 ACM Symp. on Applied Computing. ACM, Sierre, Switzerland, 2010, pp. 1584–1590

[117] Kang, X., Qin, T., Zeng, H.: ‘Countering median filtering anti-forensics and performance evaluation of forensics against intentional attacks’. Signal and Information Processing (ChinaSIP), 2015 IEEE China Summit and Int. Conf. on, Chengdu, China, 2015, pp. 483–487

[118] Mayer, O., Stamm, M.C.: ‘Countering anti-forensics of lateral chromatic aberration’. Proc. of the 5th ACM Workshop on Information Hiding and Multimedia Security, Philadelphia, Pennsylvania, USA, 2017, pp. 15–20

[119] Chen, C., Zhao, X., Stamm, M.C.: ‘Detecting anti-forensic attacks on demosaicing-based camera model identification’. IEEE Int. Conf. on Image Processing (ICIP), Beijing, China, 2017, pp. 1512–1516

[120] Cao, G., Zhao, Y., Ni, R.: ‘Forensic identification of resampling operators: a semi non-intrusive approach’, *Forensic Sci. Int.*, 2012, **216**, (1–3), pp. 29–36

[121] Peng, A., Zeng, H., Lin, X.*, et al.*: ‘Countering anti-forensics of image resampling’. IEEE Int. Conf. on Image Processing (ICIP). IEEE, Quebec City, QC, Canada, 2015, pp. 3595–3599

[122] Peng, A., Wu, Y., Kang, X.: ‘Revealing traces of image resampling and resampling antiforensics’, *Adv. Multimed.*, 2017, **2017**, pp. 1–13

[123] Li, H., Luo, W., Huang, J.: ‘Countering anti-JPEG compression forensics’. 19th Int. Conf. on Image Processing (ICIP), Orlando, FL, USA, 2012, pp. 241–244

[124] Valenzise, G., Tagliasacchi, M., Tubaro, S.: ‘Revealing the traces of JPEG compression anti-forensics’, *IEEE Trans. Inf. Forensics Sec.*, 2013, **8**, (2), pp. 335–349

[125] ‘NRCS Photo Gallary [online]’, https://photogallery.sc.egov.usda.gov/res/sites/photogallery/

[126] Wang, M., Chen, Z., Fan, W.*, et al.*: ‘Countering anti-forensics to wavelet-based compression’. IEEE Int. Conf. on Image Processing (ICIP), Paris, France, 2014, pp. 5382–5386

[127] Fahmy, G., Würtz, R.: ‘Phase based forgery detection of jpeg anti forensics’. IEEE Int. Symp. on Signal Processing and Information Technology (ISSPIT), Limassol, Cyprus, 2016, pp. 1–6

[128] Barni, M., Chen, Z., Tondi, B.: ‘Adversary-aware, data-driven detection of double JPEG compression: how to make counter-forensics harder’. IEEE Int. Workshop on Information Forensics and Security (WIFS), Abu Dhabi, United Arab Emirates, 2016, pp. 1–6

[129] Dang Nguyen, D.T., Pasquini, C., Conotter, V.*, et al.* ‘Raise: a raw images dataset for digital image forensics’. Proc. of the 6th ACM Multimedia Systems Conf. ACM, Portland, Oregon, 2015, pp. 219–224

[130] Zeng, H., Yu, J., Kang, X.*, et al.*: ‘Countering JPEG anti-forensics based on noise level estimation’, *Sci. China Inform. Sci.*, 2018, **61**, (3), p. 032103

[131] Bhardwaj, D., Pankajakshan, V.: ‘A JPEG blocking artifact detector for image forensics’, *Signal Process., Image Commun.*, 2018, **68**, pp. 155–161

[132] Bhardwaj, D., Kumawat, C., Pankajakshan, V.: ‘A method for detecting JPEG anti-forensics’. 6th National Conf. on Computer Vision, Pattern Recognition, Image Processing, and Graphics, Mandi, India, 2017, pp. 190–197

[133] Li, H., Luo, W., Qiu, X.*, et al.*: ‘Identification of various image operations using residual-based features’, *IEEE Trans. Circuits Syst. Video Technol.*, 2018, **28**, (1), pp. 31–45

[134] Li, B., Zhang, H., Luo, H.*, et al.*: ‘Detecting double JPEG compression and its related anti-forensic operations with CNN’, *Multimedia Tools Appl.*, 2019, **78**, (7), pp. 8577–8601

[135] Singh, G., Singh, K.: ‘Counter JPEG anti-forensic approach based on the second-order statistical analysis’, *IEEE Trans. Inf. Forensics Sec.*, 2019, **14**, (5), pp. 1194–1209

[136] Fahmy, G., Alqallaf, A., Wurtz, R.: ‘Phase based detection of JPEG counter forensics’. IEEE Int. Conf. on Electronics, Circuits, and Systems (ICECS), Cairo, Egypt, 2015, pp. 37–40

[137] Valenzise, G., Nobile, V., Tagliasacchi, M.*, et al.*: ‘Countering JPEG anti-forensics’. 18th IEEE Int. Conf. on Image Processing (ICIP), Brussels, Belgium, 2011, pp. 1949–1952

[138] Rudin, L.I., Osher, S., Fatemi, E.: ‘Nonlinear total variation based noise removal algorithms’, *Phys. D, Nonlinear Phenom.*, 1992, **60**, (1), pp. 259–268

[139] Fontani, M., Bonchi, A., Piva, A.*, et al.*: ‘Countering anti-forensics by means of data fusion’. Media Watermarking, Security, and Forensics Int. Society for Optics and Photonics, San Francisco, CA, 2014, vol. 9028, pp. 90280Z–90280Z–15

[140] Stamm, M.C., Liu, K.R.: ‘Forensic estimation and reconstruction of a contrast enhancement mapping’. IEEE Int. Conf. on Acoustics Speech and Signal Processing (ICASSP), Dallas, TX, USA, 2010, pp. 1698–1701

[141] Sun, J.Y., Kim, S.W., Lee, S.W.*, et al.*: ‘A novel contrast enhancement forensics based on convolutional neural networks’, *Signal Process., Image Commun.*, 2018, **63**, pp. 149–160

[142] Yang, P., Ni, R., Zhao, Y.*, et al.*: ‘Robust contrast enhancement forensics using convolutional neural networks’, arXiv preprint arXiv:180304749, 2018

[143] Shan, W., Yi, Y., Huang, R.*, et al.*: ‘Robust contrast enhancement forensics based on convolutional neural networks’, *Signal Process., Image Commun.*, 2019, **71**, pp. 138–146

[144] Akhtar, Z., Khan, E.: ‘Revealing the traces of histogram equalisation in digital images’, *IET Image Process.*, 2017, **12**, (5), pp. 760–768

[145] Zeng, H., Kang, X., Peng, A.: ‘A multi-purpose countermeasure against image anti-forensics using autoregressive model’, *Neurocomputing*, 2016, **189**, pp. 117–122

[146] Pevny, T., Bas, P., Fridrich, J.: ‘Steganalysis by subtractive pixel adjacency matrix’, *IEEE Trans. Inf. Forensics Sec.*, 2010, **5**, (2), pp. 215–224

[147] Yu, J., Zhan, Y., Yang, J.*, et al.*: ‘A multi-purpose image counter-anti-forensic method using convolutional neural networks’. Int. Workshop on Digital Watermarking. (vol. 10082 LNCS, Beijing, China, 2016), pp. 3–15

[148] Chen, Y., Kang, X., Wang, Z.J.*, et al.*: ‘Densely connected convolutional neural network for multi-purpose image forensics under anti-forensic attacks’. Proc. of the 6th ACM Workshop on Information Hiding and Multimedia Security, Innsbruck, Austria, 2018, pp. 91–96

[149] Li, H., He, P., Wang, S.*, et al.*: ‘Learning generalized deep feature representation for face anti-spoofing’, *IEEE Trans. Inf. Forensics Sec.*, 2018, **13**, (10), pp. 2639–2652