

Understanding digital image anti‑forensics: an analytical review

Neeti Taneja1 [·](http://orcid.org/0000-0002-8895-9173) Vijendra Singh Bramhe1 · Dinesh Bhardwaj2 · Ashu Taneja3

Received: 28 February 2022 / Revised: 22 February 2023 / Accepted: 15 May 2023 / Published online: 22 June 2023 © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Image forensics is essential for detecting image manipulation, authenticating images, and identifying sources of images. A forensic analyst can make use of various artifacts to develop a powerful forensic technique. These artifacts include JPEG blocking and quantization artifacts, streaking artifacts and contrast enhancement artifacts, etc. With the introduction of anti-forensics, it has become difficult for forensic experts to identify forged images. There are various anti-forensic methods available that try to eradicate these detection footprints/artifacts to fool the existing forensic detectors. Thus the detection of anti-forensic attacks is very crucial and plays a vital role in forensic analysis. This paper presents a review of various types of anti-forensic attacks, such as JPEG anti-forensics, Contrast enhancement anti-forensics, and Median fltering anti-forensics. Firstly a brief introduction is given about image forgery, JPEG compression, contrast enhancement, and median fltering. Then, anti-forensics is described in detail, and fnally, the recent state-of-the-art anti-forensic techniques are summarized in tabular form for better understanding. This may be helpful for the forensic analyst to develop robust methods for forgery detection that can be applied in various applications such as the identifcation of cybercrimes, identity thefts, etc.

Keywords Blocking artifacts · Anti-forensics · JPEG compression · MFR · CNN

 \boxtimes Neeti Taneja neeti.taneja30@gmail.com

> Vijendra Singh Bramhe vijendra.singh1@sharda.ac.in

Dinesh Bhardwaj dinesh.bhardwaj@thapar.edu

Ashu Taneja ashu.taneja@chitkara.edu.in

- ¹ Department of Computer Science & Engineering, Sharda School of Engineering & Technology, Sharda University, Greater Noida, Uttar Pradesh 201310, India
- ² Department of Electronics & Communication Engineering, Thapar Institute of Engineering and Technology, Patiala, Punjab 147004, India
- ³ Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab 140401, India

Vijendra Singh Bramhe, Dinesh Bhardwaj and Ashu Taneja contributed equally to this work.

1 Introduction

With the development of Information and Communication Technology (ICT), computers and smartphones have become essential parts of our lives. Therefore, all the data, including images, are stored in digital form. Social media platforms have become very popular for sharing information. Digital images are used to convey visual information on social media websites that spread like wildfre. But, it is easy to fabricate/manipulate a digital image with the widespread availability of powerful multimedia processing techniques and photo editing software like Adobe Photoshop, Hornil Stylepix, ACDsee, etc. Manipulated images may be used for malicious purposes, such as defaming any individual or as false evidence in a court of law. The spreading of manipulated images may even lead to communal violence or political crisis. So, it has become of paramount importance to verify the authenticity and integrity of digital images. To solve the above problem, digital image forensics comes into the picture [\[28](#page-20-0)]. Digital image forensics is a new research area that deals with multimedia security and mainly covers two things, source camera identifcation, and image forgery detection. In source camera identifcation, we have to identify which camera has taken a given photograph [[69](#page-21-0)]. In image forgery detection, the given image is checked for any kind of modifcation [\[24\]](#page-19-0). Image forensic techniques exploit the traces left by the acquisition devices or the signal-processing operations involved in the forgery. Although there exist a large number of digital forensic techniques, most of them do not account for the possibility of anti-forensics which tries to hide the signal processing traces that are exploited in forensic analysis. Authentication of digital images in the presence of anti-forensics is a challenge for multimedia security researchers. To tackle such scenarios, it is necessary to develop powerful digital forensic techniques which can detect forgery even after the application of anti-forensics or can detect the traces left by anti-forensic operations. In [\[59\]](#page-21-1), the Markov-based features extracted using discrete cosine transform (DCT) are combined with the local binary pattern (LBP) features to detect the manipulated images. Forensics doesn't only mean to deal with image forgery or identifcation of source cameras, it also refers to diferent areas like fnger vein recognition/authentication [[47](#page-20-1)], forged document detection [[13](#page-19-1)] etc. Recently various deep learning techniques have been proposed for determining the authenticity of fnger vein [[45](#page-20-2), [46](#page-20-3)]. The main objectives/contributions of this paper are as follows:

- To thoroughly review the anti-forensic techniques proposed in the literature, which are deliberately designed to fool existing forensic detectors.
- To help the reader to precisely draw a line between forensics and anti-forensics.
- To guide the forensic analyst to develop a robust image forgery detection technique in the presence of anti-forensic operations.

1.1 Types of image forgery

Image forgery is classifed into three main types:

- a) Copy-move forgery: This type of image tampering involves copying some region from a particular location in an image and pasting it at one or more places within the same image or on a different image containing the same scene $[62]$ $[62]$ $[62]$. An example $[71]$ $[71]$ of copymove forgery is shown in Fig. [1](#page-2-0):
- b) Image retouching: Such a forgery is obtained from the modifcation of the objects in terms of color or texture, intensifcation of the weather conditions, or introduction of

(a) Authentic image

blur in defusing the objects. Image retouching is illustrated with an example [[72\]](#page-21-4) as shown in Fig. [2:](#page-2-1)

c) Compositing or Image splicing: This type of forgery involves a process in which more than one image is combined to create a new tampered image. Image splicing is shown in Fig. [3](#page-3-0) where authentic image 1 and authentic image 2 are used to create a forged [[73\]](#page-21-5) image.

1.2 Image forensics

Digital image forensics aims to explore the authenticity of digital images by examining the accuracy of data and by restoring the past data of an image associated with its attainment stage [[15](#page-19-2), [49](#page-20-4)].

(a) Forged image

(b) Authentic image

(a) Authentic image1

(b) Authentic image2

 (c) Forged image

Fig. 3 Image Splicing forgery

Image forensics techniques are broadly categorized into two types:

- a) Active techniques: These include additional information that is embedded into digital images in advance. Digital watermarking is an example of an active technique where a watermark is embedded in the multimedia fle during its creation. When someone attempts to manipulate the fle, the watermark gets destroyed. This ensures some control over the manipulation of content in the multimedia fle.
- b) Passive techniques: These do not include any additional information to be embedded into digital images in advance. These techniques are also known as blind forensic techniques because they do not require prior knowledge or embedding watermarking capability.

Forensic analysis can make use of specifc footprints of various image-processing artifacts. These involve JPEG compression artifacts, Contrast enhancement, Median fltering, etc.

1.2.1 JPEG compression

In JPEG compression, the image is first divided into 8×8 sized pixels of non-overlapping blocks. Further, the processing is applied to these blocks individually. Consider an 8×8 image block $A(i, j)i$, $j \in \{0...7\}$, which is modified into two-dimensional discrete cosine transform (2D-DCT). Let $B(i, j)$ be the obtained DCT coefficient which is quantized with a quantization step size $C(i, j)$. The resultant obtained quantized DCT coefficient $D(i, j)$ is given as:

$$
D(i,j) = \left[\frac{B(i,j)}{C(i,j)}\right], \qquad i,j \in \{0, \cdots, 7\}.
$$
 (1)

where [·] denotes the integer rounding operation. Finally, entropy encoding is applied to the quantized DCT coefficients to obtain the JPEG bit-stream. A JPEG file can be translated back to the dimensional area by carrying out the inverse of the steps involved in the encoding process. The JPEG bit-stream is entropy decoded and dequantized to obtain the DCT coefficients as,

$$
B'(i,j) = D(i,j) \cdot C(i,j), \qquad i,j \in \{0, \cdots, 7\}.
$$
 (2)

The steps of JPEG compression and decompression are shown in Fig. [4](#page-4-0):

a) JPEG artifacts: There are two types of compression artifacts in JPEG forensics: Quantization artifacts and Blocking artifacts. Quantization artifacts [\[41\]](#page-20-5) involve clustering of DCT coefficients around integer multiples of the quantization step size. In the case of multiple JPEG compression experienced by an image, the periodic pattern DCT sub-band histograms are introduced with a specifc quality factor corresponding to individual compression. Blocking artifacts [[17](#page-19-3)] represent pixel discontinuities across block boundaries of the decompressed image.

A JPEG image and a zoomed-in view of the JPEG image are shown in Fig. [5](#page-5-0). It can be observed that due to the presence of blocking artifacts in JPEG images, the visual quality of the image gets degraded. The compression algorithm may be combined with encryption to provide the secured transmission/storage of data [\[35,](#page-20-6) [42\]](#page-20-7).

1.2.2 Contrast enhancement

Contrast enhancement is the process that makes the image features stand out more clearly by making optimal use of the colors available on display. The contrast of an image can be derived from its histogram. Contrast enhancement changes the pixel values in such a way that a wide range of intensity values gets covered. Various techniques, such as histogram stretching, gamma correction, etc., are used to enhance the contrast of an image. Forgers generally blend the copied area into the target image to create a forgery. The introduction of peaks and gaps in the image histogram can then be used for contrast enhancement/forgery detection.

A Contrast Enhanced (CE) image is shown in Fig. [6.](#page-6-0)

1.2.3 Median fltering

Median fltering is a non-linear technique often used to remove noise from an image. The median flter works by rolling a window over each pixel and restoring that pixel value by a median of all the neighboring pixels, including itself, specifed by the window. A forger may make use of median fltering to abolish the footprints of resampling and JPEG blocking artifacts. [[44](#page-20-8)]

Median fltering can be detected on the basis of streaking artifacts or blotching efect. The streaking artifacts represent a forensic probabilistic feature that contains the frequency

JPEG Compression and Decompression Cycle

(a) Uncompressed image

(c) zoomed-in view of JPEG image

Fig. 5 Example of JPEG blocking artifacts

occurrence of 0 and 1 bins on the histogram of frst-order pixel value diference. The value of this probabilistic feature comes out to be around 1 for original images and greater than 1 for median fltered images.

To assess the capability of the forensic detectors, there arises a requirement for antiforensic techniques. In practical terms, a fabricator tries to escape forensic detection by applying various anti-forensic techniques such as JPEG anti-forensics, Contrast enhancement anti-forensics and Median fltering anti-forensics, etc.

1.3 Anti‑forensics

It refers to concealing the traces of diferent image processing operations which are exploited by a forensic analyst for forgery detection. It is mainly introduced to fool the

(a) Uncompressed image

existing forensic detectors. Forgery is created by applying various image operations, like resizing, rotation, etc., on the copied region before pasting it to the target location in the image. It is of prime importance because the forged images are generally saved in JPEG format, and most of the forgery detection techniques exploit the JPEG compression artifacts.

The main goal of JPEG anti-forensics is to remove the diferent artifacts of JPEG compression to make the compressed image look like it has never been compressed. Therefore, the forensic analyst needs to develop an efficient and powerful detector that is robust to anti-forensic attacks.

This is illustrated in Fig. [7](#page-7-0) with the help of DCT histogram of a particular subband for an uncompressed image, JPEG image, and the image obtained after applying JPEG anti-forensic operation [\[54\]](#page-21-6). For this, the genuine uncompressed image is frst JPEG compressed with a quality factor of 60, and then the anti-forensic operation [\[54\]](#page-21-6) is applied to the resulting image. Further, the $5th$ AC subband (zig-zag order) is considered for plotting the DCT histograms where the quantization step size is 8. It can be observed that a comblike structure is obtained from the histogram of the JPEG image, and the histogram bins are centered at multiples of quantization step size 8. It is very hard to distinguish between the uncompressed image and the image obtained after applying JPEG anti-forensics as the histograms of the uncompressed image and the image obtained after applying JPEG antiforensic operation look quite similar. Table [1](#page-7-1) shows the true positive rate (TPR) values for different blocking artifact detectors [[5](#page-19-4), [17,](#page-19-3) [18](#page-19-5)] in the presence of JPEG anti-forensic operations [\[19,](#page-19-6) [54](#page-21-6)]. The false positive rate (FPR) is fxed at 0.1 for threshold calculation. It can be deduced from the table that TPR values generally decrease with an increase in quality factor. Further blocking artifact detector [[5](#page-19-4)] is the only efective detector in the presence of anti-forensic operations.

Contrast enhancement anti-forensics is illustrated in Fig. [8](#page-8-0) by considering the histograms of an original never enhanced image, normal CE image, and contrast-enhanced image obtained using anti-forensic operation [\[23\]](#page-19-7). For this, gamma correction is applied to the given image to get a normal contrast-enhanced image. Further, the anti-forensic operation [\[23\]](#page-19-7) is applied to get an anti-forensically contrast-enhanced (ACE) image. The gamma

Fig. 7 DCT histograms of (**a**) uncompressed image. (**b**) JPEG image. (**c**) image after applying JPEG antiforensic operation

value is taken as 0.6 in both cases. The peak and gap artifacts can be clearly observed in the histogram of a normal contrast-enhanced image which is obtained using gamma correction. In contrast, the histogram of the anti-forensically contrast-enhanced image appears to be similar to that of the original never enhanced image.

Fig. 8 Spatial domain histograms of (**a**) Original image. (**b**) normal CE image. (**c**) anti-forensically contrast-enhanced image

Median fltering anti-forensics explores the statistical modifcations based on the histograms of pixel value diference. The probability of zero values in the frst-order pixel value diference map increases with the application of median fltering. Anti-forensics of median fltering is proposed to fool the forensic analyst by abolishing the median fltering artifacts left by statistical changes in edge regions, blurring, etc.

A complete outlook of digital image forensic and anti-forensic scenario is depicted in Fig. [9:](#page-9-0)

2 Literature review

This section presents a detailed overview of the image anti-forensics methods proposed in the literature. For better understanding, the summary of current state-of-the-art methods is also presented in tabular form.

2.1 JPEG anti‑forensics

The literature contains a number of research papers on JPEG anti-forensics. In a pioneering work, an anti-forensic method [[55\]](#page-21-7) based on JPEG was proposed by Stamm et al. The method focuses on removing the quantization artifacts of JPEG images by

Fig. 9 A generalized outlook of digital image forensics scenario

adding an anti-forensic dither. Later on, a deblocking step was incorporated to eradicate the blocking artifacts $[54]$ $[54]$ $[54]$. The DCT coefficients in distinct AC sub-bands are assumed to be Laplacian distributed, and maximum likelihood estimation (MLE) was used for the estimation of the Laplacian parameter. For the removal of the quantization artifacts, a special noise called anti-forensic dither is added to the quantized DCT coefficients. For the removal of the blocking artifacts, a median fltering operation is applied and then fnally, Gaussian noise is added. The deviation of the Gaussian noise depends on the considered JPEG image's quality factor. This technique is shown to be efective in disguising current forensic JPEG detectors.

For deceiving detectors based on double JPEG compression, an anti-forensic method [[57\]](#page-21-8) based on Shrink and Zoom (SAZ) was introduced by Sutthiwan et al. The basic idea behind this technique is to preserve the good quality of the image by destroying the JPEG grid structure. For a given double-compressed JPEG image, the technique frst decompresses it to the spatial domain, and SAZ operation is then applied using bilinear interpolation. Lastly, the resultant JPEG image is compressed with the secondary quality factor. A universal anti-forensic technique was proposed by Barni et al. [[3\]](#page-19-8), which is competent for fooling multiple-JPEG compression detectors based on frst-order statistics. This method involves the revision of DCT subband histograms of multiplecompressed images in such a way that the resulting histograms coincide with those of single-compressed images.

In [[18\]](#page-19-5), Fan et al. introduced a total variation-based technique to remove blocking artifacts from a JPEG compressed image which helps in defeating diferent JPEG forensic detectors. Further, it has been shown that the calibration-based detector can be defeated by feature value optimization. A JPEG-based anti-forensic algorithm was proposed by Fan et al. [[19\]](#page-19-6) with the objective of refning the conceptual quality of the antiforensically modifed image while maintaining forensic undetectability. The algorithm involves four main steps: deblocking based on frst-round total variation (TV), adding conceptual dither in the DCT domain, deblocking based on second-round TV, and lastly, the decalibration.

An enhanced JPEG-based anti-forensic approach for removing the blocking artifacts was highlighted in [[51](#page-20-9)]. These artifacts were left through JPEG compression in both the spatial and DCT domains. The grainy noise obtained by conceptual DCT histogram smoothing

was reduced by the introduced denoising operations. Two kinds of denoising operations were used: one with a forced minimization problem of the total change of energy and a second with an adjusted weighted function. Also, for removing the blocking artifacts in the spatial domain, an enhanced TV-based deblocking was used.

The detectors based on the FSD distribution of the DCT coefficients were targeted in the method proposed in [\[43\]](#page-20-10). This method is shown to be efective in restoring the FSD distribution of either a single/double compressed image to that of the uncompressed image or a double compressed image to that of a single compressed image.

In [[4](#page-19-9)], Barni et al. provide an overview of adversarial multimedia forensics. The authors frst review the techniques proposed by a forensic analyst and a forger (attacker) independently. An attacker may create a forgery which can further introduce some other artifacts in the forged image. A forensic analyst may exploit these artifacts to locate and detect the modifed regions in a given image. An attacker may further try to create a realistic image forgery along with the suppression of artifacts and thus results in a thief-and-police game between the fabricator and the forensic expert.

A novel anti-forensic technique was introduced by [\[30\]](#page-20-11) to check the performance of existing JPEG forensic detectors. The technique comprises two main steps: Firstly, the block-shifted DCT approach is used on the input image, which is JPEG compressed to load the deviations in the arrangement of DCT coefficients. This leads to the addition of dithering noise in the image and a reduction of cost in creating forgery on those images. Secondly, the outcome of block shifted DCT approach is further modifed by using deblocking function based on TV to eradicate the leftover JPEG-blocking artifacts.

An efficient method for JPEG anti-forensics based on CNN $[1]$ $[1]$ was proposed for deceiving both JPEG and DJPEG detectors with higher-quality images so as to enhance the performance of existing anti-forensics methods. It was also demonstrated that JPEG and DJPEG detectors could be deceived by JPEG detection in the presence of current anti-forensics but fail to work well in non-aligned cases, resulting in images with lower quality.

The introduction of anti-forensic techniques prompted research on the methods for countering anti-forensics. The methods proposed by Lai et al. [[34](#page-20-12)] and Valenzise et al. [[58](#page-21-9)] are used for countering the anti-forensic technique proposed by Stamm et al. [[55](#page-21-7)]. Two discrete detectors for anti-forensically modifed images are proposed in [[34](#page-20-12)]. The frst detector makes use of the fact that the high-frequency DCT subbands are not altered by the antiforensic operation. The second detector utilizes the divergence between the deviation in DCT coefficients of a given image and that of its calibrated version. Valenzise et al. have introduced grainy noise in the spatial domain by the addition of anti-forensic dither in the DCT domain. A technique was proposed for the detection of this anti-forensic operation by using TV to measure the noisiness in a given image.

In [[36](#page-20-13)], Haodong Li et al. introduced a technique based on machine learning for the detection of anti-forensically modifed images. The authors observed that the dithering operation destroys the intra and inter-block correlation of an image. From the above surveillance, Markov random processes and transition probability matrices were used for the extraction of a 100-dimensional feature vector. For the training and classifcation of original, JPEG, and anti-forensically modifed images, the SVM classifer is used. In [\[37\]](#page-20-14), Haodong Li et al. presented a technique based on the residual domain for the detection and classifcation of various image processing operations as well as anti-forensic operations. The image residual is obtained by performing a high-pass fltering operation to crush the consequence of image details. The authors used spatial rich model (SRM) features [[21](#page-19-11)] with reduced dimensionality for the training of a multi-class SVM classifer.

The experimental results indicate an efective performance of the detector as the method outperforms when working with images tacted with JPEG-aligned and non-aligned antiforensic techniques.

A Convolutional Neural Network (CNN) based approach [[22](#page-19-12)] was proposed for the classifcation of multiple JPEG compressed images. Based on DCT histograms of luminance and chrominance color planes, signifcant features were extracted. A single feature vector was created by concatenating each histogram consisting of 21 low-frequency DCT subbands placed in zig-zag scanning order. The above-generated feature vector can be used for the training of CNN to classify multiple-JPEG compressed images. A forensic technique to counter JPEG anti-forensic attacks was proposed by [[52](#page-21-10)] for analyzing the secondorder statistics based on Co-occurrence Matrices (CMs). Three main steps were involved in the proposed method: Initially, a target diference image is selected, then CMs are evaluated, and fnally, statistical analysis for second-order is generated based on CMs. Also, this method particularly targets anti-forensic JPEG dithering as compared to other existing techniques. Thus, this method will barely generate false positives.

A novel and robust technique was proposed by [\[32\]](#page-20-15) for the detection of JPEG compressed images saved in uncompressed format. The detector mainly uses the distinct values in the arrangement of DCT coefficients in the AC subbands of the original uncompressed images and JPEG-U images. Based on the computed diference, a statistic is obtained, which is then used to compare with a threshold for the identifcation of JPEG-U images. A method for countering the JPEG anti-forensic attacks was presented by [[31](#page-20-16)]. The capability of the detector depends on the identifcation of JPEG anti-forensics. The proposed technique involves three main steps: Firstly, a target diference image is picked out, secondly, for inter and intra-block computations, Markov Transition Probability Matrices (MTPMs) are evaluated in the DCT domain, and second-order statistical features are generated based on the evaluated MTPMs, lastly, for training and classifcation of resultant features, SVM classifer is used. From the experimental results, it can be observed that the proposed forensic technique outperforms other existing techniques in terms of minimum decision error, whose small values will indicate better detection.

2.2 Contrast enhancement anti‑forensics

Various forensic methods for the detection of contrast enhancement were proposed, and these methods rely on checking image histograms for the presence of peaks and gaps. A novel forensic method was proposed for exposing cut-and-paste image forgery [\[38\]](#page-20-17) through the detection of contrast enhancement. This involves revealing the inter-channel correlation introduced by interpolated images and showing how a linear or non-linear contrast enhancement can change this natural inter-channel correlation. To measure this inter-channel correlation, a metric was used, which also helps in diferentiating the original image from a CE image.

Further, two methods were introduced by [\[39\]](#page-20-18) for the detection of contrast enhancement. The frst method uses a quadratic weighing function to measure the distortion in the image histogram resulting from the presence of contrast enhancement. While the second method uses the strategy of linear threshold to get along with the selection of the threshold. Both methods result in the efectiveness of the performance in terms of ROC curves.

Contrast enhancement creates a realistic composite image by adjusting the brightness and contrast of the image. Therefore there is a need to identify contrast enhancement in images so as to check the authenticity of digital images. Two novel techniques were

introduced for the identifcation of contrast enhancement-related manipulation [\[8](#page-19-13)]. The frst method focuses on analyzing peak/gap artifacts in the histogram introduced by JPEG compression and pixel value mappings for the identifcation of global contrast enhancement and distinguished by identifying the zero-height gap fngerprints. In the second algorithm, the clustering of block-wise peak/gap bins was done to recognize the contrast enhancement mappings applied to digital images.

Thereafter anti-forensic methods were introduced based on peak-gap artifacts [[6](#page-19-14)] of the pixel gray level histogram. After that, an alternative method [\[33\]](#page-20-19) was proposed, which tries to remove these peaks and gaps by applying random dither, thereby impacting the accuracy of the contrast enhancement detection method and, at the same time, maintaining good image quality. An anti-forensic technique based on contrast enhancement [\[23\]](#page-19-7) was proposed, which results in similar characteristics as that of the original unenhanced image and lower TV. This is realized by having negligible distortion in the frst and second-order statistics of the processed image and thus can easily fool the existing contrast enhancement detectors.

A novel counter-forensic contrast enhancement technique was proposed to suppress the captured artifacts [\[2](#page-19-15)] in the process of contrast enhancement operation. The method generates a CE image by using the information in both spatial and DCT domains. It shows robustness against various contrast enhancement detectors and efectiveness in terms of image quality metrics.

Contrast enhancement anti-forensic methods are good at removing forensic footprints from the histogram of the CE image by showing their forging capability, but at the same time, they ignore the problem of exposing changes in pixel values in the pixel domain. This problem was solved by an anti-forensic contrast enhancement method [[70](#page-21-11)] based on Generative Adversarial Network (GAN). GAN is used for the processing of CE images in the pixel domain and making it difcult to diferentiate them from the original image. A histogram-based loss is used to increase the efficiency of the attacks in both the histogram domain and the Gray Level Co-occurrence Matrix (GLCM) domain. The method shows better anti-forensic attack performance with improved image quality.

Usually, frst-order statistics obtained from histograms were used for the detection of contrast enhancement in digital forensic analysis, but methods using this scheme somewhat perform badly in the company of various counter-forensic attacks, therefore a novel forensic second-order statistics-based technique [[16](#page-19-16)] obtained through CM was presented.

An iterative algorithm for estimating contrast enhancement was introduced by [\[60\]](#page-21-12) to expose forgeries in digital image. The method recovers the diferences in the original image pixel histogram and contrast enhancement image pixel histogram so as to identify composite images. It works quite efectively in the presence of dithering noise, which can be used for locking up the footprints of contrast enhancement. A novel forensic technique for contrast enhancement based on CNN was proposed by [[56](#page-21-13)]. In this method, CNN appropriately detects modifcations and extracts suitable features by learning representations of hierarchical features and optimizing the results of classifcation. Also, CNN is fed with a GLCM, which makes it superior in terms of detecting forgeries in the company of various counter-forensic attacks.

The above-proposed methods for contrast enhancement were not efective in the presence of some image processing operations, so contrast enhancement forensics was analyzed in the pixel domain and histogram domain. Two robust methods based on CNN [[65](#page-21-14)] were introduced for the detection of contrast enhancement in the presence of JPEG compression and various anti-forensic attacks. Further, a modifed CNN-based JPEG robust method [[48](#page-20-20)] was presented for the detection of contrast enhancement. The modifed CNN accepts

enhanced images as inputs and outputs the GLCM, which contains contrast enhancement fngerprints.

Further, the above problem was solved by introducing a new deep learning structure based on a dual-domain fusion CNN [[63](#page-21-15)], which shows robustness against various contrast enhancement levels, anti-forensic attacks, and pre-JPEG compression. Contrast enhancement forensics was performed by fusing the features of the pixel domain and histogram domain. P-CNN was used to capture the pixel domain patterns, while H-CNN was responsible for the extraction of the signifcant features in the histogram domain. The fully connected layers of CNN were fed with the fused feature patterns of both domains for classifcation purposes.

2.3 Median fltering anti‑forensics

A forensic environment in the digital era requires the establishment of efficient methods for tracing out the signifcant modifcations that occurred because of the presence of median fltering. Median fltering causes the removal of forensic artifacts. A passive method [[7](#page-19-17)] was proposed for spotting the presence of median fltering, which mainly involves analyzing statistical characteristics and measuring the probability of zero values on the diference map. Further, another passive method [\[9\]](#page-19-18) was introduced based on the view that the images undergone median fltering involve distinct footprints around edges such as preservation of edges, suppression of noise, correlation of neighboring pixels, etc. Based on these median fltering traces, an edge-based prediction matrix (EBPM) is formed that carries approximate projection coefficient values among edges, and finally, the SVM classifier is used for classifcation. The above method was further improved by detecting median fltering in arbitrary images, even those images which contain very less resolution [\[66\]](#page-21-16) and are compressed using JPEG compression.

The method in [[9](#page-19-18)] was improvised by introducing another statistical technique [[10](#page-19-19)] that uses statistical features in the diferent domains of images. In this technique, the cumulative distribution function of frst-order diferences was derived, and the behavior of adjacent diference pairs was analyzed in the diference domain for original images, median fltered images, etc. There is an improvement in the method presented in [\[66](#page-21-16)] as now a novel forensic trace called Median Filter Residual(MFR) was introduced for the identifcation of median fltering from an image compressed using JPEG compression. It reduces the image edge and texture interference, which was the major limitation of the previous method. A vigorous technique for the detection of median fltering was proposed for analyzing the statistical characteristics of MFR. An autoregressive (AR) model [[26\]](#page-20-21) was used for gathering the statistical characteristics of MFR, and AR coefficients are considered as features that are used for the training of SVM for the identifcation of median fltering. The method works efectively for JPEGcompressed images with lower quality factors and shows increased performance in the presence of low false-positive rates.

Another improvement in the above median filtering detection methods had been made using an approach based on difference domain where two new feature sets [\[11](#page-19-20)] were introduced for differentiating an original uncompressed image from an image that had undergone median filtering or an image possessing average median filtering. This method shows robustness in the presence of noise and efficacy in the case of less resolution and more JPEG post-compression. Another novel method [[68](#page-21-17)] was proposed for the identification of median filtering, which requires the use of a local texture operator called local ternary pattern (LTP). The variations in the local texture were encoded by a suitable coding function, and the changes caused by median filtering in the local texture were effectively captured by LTP. The method was cost-effective by using kernel principal component analysis (KPCA) so as to reduce the feature set's dimensionality.

After that, a deep learning-based method [\[12\]](#page-19-21) for median fltering detection was proposed in which a CNN impulsively learns the hierarchical traits and performs classifcation. The method was efficient for small and JPEG-compressed image blocks and best suited for the identifcation of cut-and-paste forgeries. A frequency-domain feature-based method [[40](#page-20-22)] was introduced for distinguishing median fltered images from original, gaussian low pass fltered, and average fltered ones, signifcantly in case of less resolution and images compressed using JPEG compression. The technique shows reliability by reducing the computational time required for classifcation, thereby depicting the applicability in the real-time operation of massive interactive media.

There is improvement proposed in the method [[26](#page-20-21)] for the detection of median fltering in heavily compressed images by using a two-dimensional auto-regressive model (2D-AR) for capturing the statistical characteristics of MFR, average fltered residual (AFR), and gaussian fltered residual (GFR) respectively, and combining the 2D-autoregressive coefficients $[64]$ $[64]$ $[64]$ of the three residuals so as to obtain a feature set. The derived feature set is fed to the SVM for training and classifcation. An improved approach for the identifcation of median fltering in heavily compressed and lowresolution images was introduced for the construction of a feature set based on the residuals by including the Markov chain with the AR model [[44\]](#page-20-8). These models help in revealing dissimilar correlations among neighboring residuals, and various methods to reduce dimensionality are hired for fast detection. Anti-forensics are applied to fool the existing forensic detection of median fltering. One such anti-forensic technique [[61](#page-21-19)] was introduced by adding noise to the pixel diference distribution of images. Both the original image and the median fltered image's pixel diference distribution were estimated, and when noise distribution was added, the pixel diference arrangement of an anti-forensically modifed image seemed to surface from the original uncompressed image. To limit distortion that arises in the process of modifcation, several measures were employed.

A novel anti-forensic method [[14](#page-19-22)] was proposed to hide the traces of median fltering by using the information of features via suitable random pixel modifcation. The method shows efficacy in terms of improvement in image quality.

Further, a variational deconvolution framework [\[20\]](#page-19-23) was proposed for quality improvement and anti-forensics of the median fltered image, which consists of convolution term, fdelity term, and prior term. The estimation of the median fltering operation was done by a convolution kernel. The fdelity term keeps some image processing footprints and denoising hiding efects in processed images close to the image, which is median fltered. Finally, the prior term reconciles the pixel value outcomes of the altered image to coincide with the distribution of the original image.

An efficient anti-forensic technique $[50]$ $[50]$ $[50]$ was proposed to hide the traces of median filtering operation by ensuring a minor change in spatial characteristics. The method uses statistics of the image as a diference between Anistrophic and Isotrophic TV regularization. Another median fltering anti-forensic method [\[29\]](#page-20-24) based on CNN was introduced to eradicate the traces from median-fltered images. The GAN framework was used for the generation of images that obey the key statistical properties of the original unaltered image, thus improving forensic undetectability.

Now a two-step median fltering anti-forensics framework was introduced [[53\]](#page-21-20), which hides the median fltering traces to fool the median fltering detectors. The framework involves the generation of median-fltered forgery using a variational deconvolution approach. After that, the remains of median fltering left during the deconvolution process was removed by using TV based minimization algorithm. The method shows efectiveness on the basis of obtaining a better quality image and achieving good forensic undetectability. An improved technique better than the above methods was introduced to hide the traces of median fltering anti-forensics [\[25\]](#page-20-25), thereby reducing forensic detectability. The blurred edges in the output of median fltering were restored by Unsharp Masking(UM). The peak signal-to-noise ratio (PSNR) between the initial loaded image and the refned output image was used as the minimum error sense criterion for the evaluation of the optimum amount, which controls the degree of sharpening. The values of PSNR seem to be higher in this approach as compared to previous anti-forensic techniques such as CNN, VD, and VD with TV minimization.

For countering median fltering anti-forensics, a novel technique [[67](#page-21-21)] was introduced to analyze the fngerprints left by median fltering anti-forensics at a computationally low cost. Further, an improved counter anti-forensic technique [\[27\]](#page-20-26) was presented for the detection of anti-forensic attacks proposed in [\[67\]](#page-21-21). The technique works by adding noise adaptively in the pixel domain based on a constant signal-to-noise ratio (SNR). For better understanding, the current state-of-the-art anti-forensic detection methods are summarized in Table [2:](#page-16-0)

3 Conclusion and future scope

This paper presents a detailed review of various anti-forensic techniques used in digital image forensics. It has been shown that a forger can use various anti-forensic operations to hide the evidence of image manipulation, which can lead to false decisions in a court of law. So it is the need of the hour to stop these forgers and their wrong intentions. The main emphasis of the survey is on JPEG anti-forensics, Contrast enhancement anti-forensics, and Median fltering anti-forensics. JPEG anti-forensics involve various techniques which are capable of removing quantization and blocking artifacts and resulting in an anti-forensically modifed uncompressed image that looks similar to the original uncompressed image. So it becomes difficult for forensic experts to distinguish authentic uncompressed images and the anti-forensically modifed uncompressed image. The median fltering and contrast enhancement operations play an important role in forensic analysis. The methods for contrast enhancement and median fltering anti-forensics are elaborated along with the JPEG anti-forensics. Currently, most of the existing forensic detectors either get fooled by the various anti-forensic attacks or do not account for the possibility of anti-forensics. The presented survey may be utilized by a forensic expert to design robust techniques that will be able to handle the abovementioned problems and achieves better accuracy. Further, forensic experts can take ideas from the presented study to combine two existing techniques to detect diferent anti-forensic operations. It is possible that these techniques may not perform well if applied individually as compared to when they are combined together. The limitation of the presented study is that it does not include the various aspects of video forensics/ anti-forensics, which will be included in future work.

 \hat{Z} Springer

Declarations

The authors declare that they do not have any personal or fnancial confict of interests.

References

- 1. Kim D, Ahn W, Lee HK (2021) End to end anti-forensics method of single and double JPEG detection. IEEE Access 9:13390–13402
- 2. Mehrish A, Subramanyam AV, Emmanuel S (2019) Joint spatial and discrete cosine transform domain based counter forensics for adaptive contrast enhancement. IEEE Access 7:27183–27195
- 3. Barni M, Fontani M, Tondi B (2015) Universal counterforensics of multiple compressed JPEG images. In: Proceedings of the 13th International Workshop Digital Forensics and Watermarking, p 31–46
- 4. Bayar B, Stamm MC (2018) Adversarial multimedia forensics: Overview and challenges ahead. In: Proceedings of the European Signal Processing Conference (EUSIPCO)
- 5. Bhardwaj D, Pankajakshan V (2018) A JPEG blocking artifact detector for image forensics. Signal Process Image Commun 68:155–161
- 6. Cao G, Zhao Y, Ni R, Tian H (2010) Anti-forensics of contrast enhancement in digital images. In: Proceedings of the 12th ACM workshop on Multimedia and security, p 25–34
- 7. Cao G, Zhao Y, Ni R, Yu L, Tian H (2010) Forensic detection of median fltering in digital images. In: Proceedings of the IEEE international conference on multimedia expo, Singapore, p 89–94
- 8. Cao G, Zhao Y, Ni R, Li X (2014) Contrast enhancement-based forensics in digital images. IEEE Trans Inf Forensic Secur 9:515–525
- 9. Chen C, Ni J (2011) Median fltering detection using edge based prediction matrix. In: Proceedings of the 10th international workshop on digital forensics and watermarking, p 361–375
- 10. Chen C, Ni J, Huang R, Huang J (2012) Blind median fltering detection using statistics in diference domain. In: Information Hiding, IH 2012, volume 7692, p 1–15
- 11. Chen C, Ni J, Huang J (2013) Blind detection of median fltering in digital images: a diference domain based approach. IEEE Trans Image Process 22:4699–4710
- 12. Chen J, Kang X, Liu Y, Wang ZJ (2015) Median fltering forensics based on convolutional neural networks. IEEE Signal Process Lett 22(11):1849–1853
- 13. Cruz F, Sidere N, Coustaty M, D'Andecy VP, Ogier JM (2017) Local binary patterns for document forgery detection. In: Proceedings of the 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, volume 01, p 1223–1228
- 14. Dang-Nguyen DT, Gebru ID, Conotter V, Boato G, De Natale FGB (2013) Counter-forensics of median fltering. In: Proceedings of the IEEE 15th International Workshop on Multimedia Signal Processing (MMSP), Pula, Italy, p 260–265
- 15. Datta P, Rani S, Koundal D (2020) Detection of eye ailments using segmentation of blood vessels from eye fundus image. In: Proceedings of the ICRIC 2019, p 515–531
- 16. De Rosa A, Fontani M, Massai M, Piva A, Barni M (2015) Second-order statistics analysis to cope with contrast enhancement counter-forensics. IEEE Signal Process Lett 22(8):1132–1136
- 17. Fan Z, De Queiroz RL (2003) Identifcation of bitmap compression history: JPEG detection and quantizer estimation. IEEE Trans Image Process 12(2):230–235
- 18. Fan W, Wang K, Cayre F, Xiong Z (2013) A variational approach to JPEG anti-forensics. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2013, Vancouver, p 3058–3062
- 19. Fan W, Wang K, Cayre F, Xiong Z (2014) JPEG anti-forensics with improved tradeof between forensic undetectability and image quality. IEEE Trans Inf Forensic Secur 9(8):1211–1226
- 20. Fan W, Wang K, Cayre F, Xiong Z (2015) Median fltered image quality enhancement and anti-forensics via variational deconvolution. IEEE Trans Inf Forensic Secur 10(5):1076–1091
- 21. Fridrich J, Kodovsky J (2012) Rich models for steganalysis of digital images. IEEE Trans Inf Forensic Secur 7(3):868–882
- 22. Priyanka, Singh G, Singh K(2020) An improved block based copy-move forgery detection technique. Multimed Tools Appl 79(5):13011–13035
- 23. Ravi H, Subramanyam AV, Emmanuel S (2015) Ace - an efective anti-forensic contrast enhancement technique. IEEE Signal Process Lett 23(2):212–216
- 24. Hingrajiya KH, Sheth RK (2021) Comparative study of digital image forgery detection techniques. In: Proceedings of the International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), p 83–86
- 25. Kaimal A, Priestly B (2019) Removing the traces of median fltering via unsharp masking as an antiforensic approach in medical imaging. Biomed Pharmacol J 12(3):1395–1402
- 26. Kang X, Stamm MC, Peng A, Ray Liu KJ (2013) Robust median fltering forensics using an autoregressive model. IEEE Trans Inf Forensic Secur 8:1456–1468
- 27. Kang X, Qin T, Zeng H (2015) Countering median fltering anti-forensics and performance evaluation of forensics against intentional attacks. In: Proceedings of the IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP), p 483–487
- 28. Kejun Zhang Y, Liang JZ, Wang Z, Li X (2019) No one can escape: a general approach to detect tampered and generated image. IEEE Access 7:129494–129503
- 29. Kim D, Jang HU, Mun SM, Choi S, Lee HK (2018) Median fltered image restoration and anti-forensics using adversarial networks. IEEE Signal Process Lett 25(2):278–282
- 30. Kumar A, Kansal A, Singh K (2019) An improved anti-forensic technique for JPEG compression. Multimed Tools Appl 78(18):25427–25453
- 31. Kumar A, Singh G, Kansal A, Singh K (2021) Digital image forensic approach to counter the JPEG anti-forensic attacks. IEEE Access 9:4364–4375
- 32. Kumawat C, Pankajakshan V (2020) A robust JPEG compression detector for image forensics. Signal Process Image Commun 89:116008
- 33. Kwok CW, Oscar A, Chui SH (2011) Alternative anti-forensics method for contrast enhancement. In: Proceedings of the 10th International Workshop, IWDW 2011, p 398–410
- 34. Lai S, Böhme R (2011) Countering counter-forensics: the case of JPEG compression. Information hiding. Lecture notes in computer science, Springer, 6958:285–298
- 35. Lalitha RVSS, Naga Srinivasu P (2017) An efficient data encryption through image via prime order symmetric key and bit shuffle technique. In: proceedings of the computer communication, networking and internet security, Springer Singapore, Singapore, p 261–270
- 36. Li H, Luo W, Huang J (2012) Countering anti-JPEG compression forensics. In: Proceedings of the 19th IEEE International Conference on Image Processing, p 241–244
- 37. Li H, Luo W, Qiu X, Huang J (2018) Identifcation of various image operations using residual-based features. IEEE Trans Circ Syst Vid Technol 28(1):31–45
- 38. Lin X, Li CT, Hu Y (2013) Exposing image forgery through the detection of contrast enhancement. In: Proceedings of the IEEE International Conference on Image Processing, p 4467–4471
- 39. Lin X, Wei X, Li CT (2014) Two improved forensic methods of detecting contrast enhancement in digital images. In: Proceedings of SPIE - The International Society for Optical Engineering 9028
- 40. Liu A, Zhao Z, Zhang C, Yuting S (2017) Median fltering forensics in digital images based on frequency-domain features. Multimed Tools Appl 76
- 41. Luo W, Huang J, Qiu G (2010) JPEG error analysis and its applications to digital image forensics. IEEE Trans Inf Forensic Secur 5(3):480–491
- 42. Naga Srinivasu P, Norwawi N, Amiripalli SS, Deepalakshmi P (2022) Secured compression for 2d medical images through the manifold and fuzzy trapezoidal correlation function. Gazi Univ J Sci 35(4):1372–1391
- 43. Pasquini C, Comesana-Alfaro P, Pérez-González F, Boato G (2014) Transportation-theoretic image counterforensics to frst signifcant digit histogram forensics. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), p 2699–2703
- 44. Peng A, Luo S, Zeng H, Yadong W (2019) Median fltering forensics using multiple models in residual domain. IEEE Access 7:28525–28538
- 45. Shaheed K, Mao A, Qureshi I, Abbas Q, Kumar M, Zhang X (2022) Finger-vein presentation attack detection using depthwise separable convolution neural network. Expert Syst Appl 198:116786
- 46. Shaheed K, Mao A, Qureshi I, Kumar M, Hussain S, Ullah I, Zhang X (2022) Ds-cnn: a pre-trained xception model based on depth-wise separable convolutional neural network for fnger vein recognition. Expert Syst Appl 191:116288
- 47. Shaheed K, Mao A, Qureshi I, Kumar M, Hussain S, Zhang X (2022) Recent advancements in fnger vein recognition technology: methodology, challenges and opportunities. Inf Fus 79:84–109
- 48. Shan RHW, Yi Y, Xie Y (2019) Robust contrast enhancement forensics based on convolutional neural networks. Signal Process Image Commun 71:138–146
- 49. Sharma K, Singh G, Goyal P (2023) Ipdcn2: Improvised patch-based deep cnn for facial retouching detection. Expert Syst Appl 211:118612
- 50. Sharma S, Subramanyam AV, Jain M, Mehrish A, Emmanuel S (2016) Anti-forensic technique for median fltering using L1- L2 TV model. In: Proceedings of the IEEE International Workshop on Information Forensics and Security (WIFS), p 1–6
- 51. Singh G, Singh K (2017) Improved JPEG anti-forensics with better image visual quality and forensic undetectability. Forensic Sci Int 277:133–147
- 52. Singh G, Singh K (2019) Counter JPEG anti-forensic approach based on the second-order statistical analysis. IEEE Trans Inf Forensic Secur 14(5):1194–1209
- 53. Singh K, Kansal A, Singh G (2019) An improved median fltering anti-forensics with better image quality and forensic undetectability. Multidim Syst Sign Process, p 1–24
- 54. Stamm MC, Ray Liu KJ (2011) Anti-forensics of digital image compression. IEEE Trans Inf Forensic Secur 6(3):1050–1065
- 55. Stamm MC, Tjoa SK, Lin WS, Liu KJR (2010) Anti-forensics of JPEG compression. In: Proceedings of the IEEE International Conference Acoustics, Speech and Signal Processing, Dallas, p 1694–1697
- 56. Sun JY, Kim SW, Lee SW, Ko SJ (2018) A novel contrast enhancement forensics based on convolutional neural networks. Signal Process Image Commun 63:149-160
- 57. Sutthiwan P, Shi YQ (2012) Anti-forensics of double JPEG compression detection. In: Proceedings of the 10th International Workshop Digital Forensics and Watermarking, Springer, Berlin, Heidelberg, p 411–424
- 58. Valenzise G, Tagliasacchi M, Tubaro S (2013) Revealing the traces of JPEG compression anti-forensics. IEEE Trans Inf Forensic Secur 8(2):335–349
- 59. Walia S, Kumar K, Kumar M, Gao XZ (2021) Fusion of handcrafted and deep features for forgery detection in digital images. IEEE Access 9:99742–99755
- 60. Wen L, Qi H, Lyu S (2018) Contrast enhancement estimation for digital image forensics. ACM Trans Multimed Comput Commun Appl (TOMM) 14(2):1–21
- 61. Wu ZH, Stamm MC, Liu KJR (2013) Anti-forensics of median fltering. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, p 3043–3047
- 62. Xie H, Ni J, Shi YQ (2021) Dual-domain generative adversarial network for digital image operation anti-forensics. IEEE Trans Circ Syst Vid Technol 32(3):1701–1706
- 63. Yang P (2021) Dual-domain fusion convolutional neural network for contrast enhancement forensics. Entropy 23:1318
- 64. Yang J, Ren H, Zhu G, Huang J, Shi YQ (2017) Detecting median fltering via two-dimensional AR models of multiple fltered residuals. Multimed Tools Appl 77:7931–7953
- 65. Yang P, Ni R, Zhao Y, Zhao W (2018) Robust contrast enhancement forensics using pixel and histogram domain CNNs. 03
- 66. Yuan HD (2011) Blind forensics of median fltering in digital images. IEEE Trans Inf Forensic Secur 6(4):1335–1345
- 67. Zeng H, Qin T, Kang X, Liu L (2014) Countering anti-forensics of median fltering. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),Florence, p 2704–2708
- 68. Zhang Y, Li S, Wang S, Shi YQ (2014) Revealing the traces of median fltering using high-order local ternary patterns. IEEE Signal Process Lett 21:275–279
- 69. Zhao M, Wang B, Wei F, Zhu M, Sui X (2020) Source camera identifcation based on coupling coding and adaptive flter. IEEE Access 8:54431–54440
- 70. Zou H, Pengpeng Y, Ni R, Zhao Y (2021) Anti-forensics of image contrast enhancement based on generative adversarial network. Secur Commun Netw 2021:1–8
- 71. Tralic D, Zupancic I, Grgic S, Grgic M (2013) CoMoFoD - New Database for Copy-Move Forgery Detection. In: Proceedings of the 55th International Symposium ELMAR-2013, p 49–54
- 72. Qazi T, Hayat K, Khan SU, Madani SV, Khan IA, Kolodziej Joanna, Hongxiang L, Weiyao Lin, Choong Yow Kin, Cheng-Zhong Xu (2013) Survey on blind image forgery detection. IET Image Processing 7(7):660–670
- 73. Dong J, Wang W, Tan T (2013) CASIA Image Tampering Detection Evaluation Database. In: Proceedings of the IEEE China Summit and International Conference on Signal and Information Processing, Beijing, China, p 422–426

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.