Unravelling Digital Forgeries: A Systematic Survey on Image Manipulation Detection and Localization

VIJAYAKUMAR KADHA, National Institute of Technology Rourkela, Rourkela, India and Amrita Vishwa Vidyapeetham, Amaravati, India
SAMBIT BAKSHI, Computer Science and Engineering, National Institute of Technology Rourkela, Rourkela, India
SANTOS KUMAR DAS, Electronics and Communication Engineering, National Institute of Technology Rourkela, Rourkela, India

In recent years, deep learning has made significant strides, especially in computer vision applications and, more specifically, in information forensics. On the other hand, data-driven approaches have shown much promise in identifying manipulations in images and videos. However, most forensic tools ignore deep learning in favour of more traditional methodologies. This article thoroughly analyses the current state-of-the-art methods for detecting and localizing image alteration using classical and deep learning-based algorithms. In addition, this review includes the latest developments in the digital image forensics field, including Convolutional Neural Networks (CNNs), while incorporating insights from classical approaches and machine learning models. Furthermore, the most significant data-driven techniques to address the issue of image manipulation detection and localization are presented and segregated into four subtopics: copy-move, splicing, object removal, and contrast enhancement. This study provides an exhaustive and up-to-date survey of the field for researchers and practitioners working in this domain. In addition, it covers the current challenges and future directions in deep learning for image manipulation detection and localization. Finally, this review's discussion of relevant approaches and experiments will aid future exploration and development in this field.

CCS Concepts: • **Computing methodologies** → **Image manipulation**; **Image representations**; **Machine learning**; • **Applied computing** → **Investigation techniques;**

Additional Key Words and Phrases: Image manipulation, investigation techniques, image representations, machine learning

ACM Reference Format:
VijayaKumar Kadha, Sambit Bakshi, and Santos Kumar Das. 2025. Unravelling Digital Forgeries: A Systematic Survey on Image Manipulation Detection and Localization. *ACM Comput. Surv.* 57, 12, Article 323 (July 2025), 36 pages. https://doi.org/10.1145/3731243

Authors' Contact Information: VijayaKumar Kadha, National Institute of Technology Rourkela, Rourkela, Odisha, India and Amrita Vishwa Vidyapeetham, Amaravati, Andhra Pradesh, India; e-mail: k_vijayakumar@av.amrita.edu; Sambit Bakshi, Computer Science and Engineering, National Institute of Technology Rourkela, Rourkela, Odisha, India; e-mail: sambitbaksi@gmail.com; Santos Kumar Das, Electronics and Communication Engineering, National Institute of Technology Rourkela, Rourkela, Odisha, India; e-mail: dassk@nitrkl.ac.in.

This work is licensed under a Creative Commons Attribution 4.0 International License.

© 2025 Copyright held by the owner/author(s).
ACM 0360-0300/2025/07-ART323
https://doi.org/10.1145/3731243

1 Introduction

Multimedia data is gradually becoming permissible as evidence in a court of law, and there is a requirement to identify the life cycle of the image under investigation in several forensic applications. However, image alteration is now feasible with the advancement of affordable photo-editing software and instantaneously possible since the spectacular development of computer and network technology. For better understanding, some common image manipulations are presented in Table 1. Due to this, numerous altered photos have been widely disseminated in our daily lives through social media platforms. Over the last two decades, the academic community and law enforcement have paid growing attention towards image forensic technology to restore digital image reliability. Therefore, solid and dependable forensic techniques for detecting image manipulations require inspecting the image's information, lighting, and colouration, among other details, for evidence of tampering.

On the other hand, with the rise of digital media and straightforward access to image editing tools, image manipulation has become a widespread problem, particularly in fake news, propaganda, and digital forensics. For instance, various image manipulation techniques like Face Swap, Deepfake, and computer-generated images are illustrated in Figure 1. The first row demonstrates the Face Swap process, where the initial two images represent the original input faces, and the third image depicts the resultant face-swapped output. The second row showcases Deepfake images generated by the source inputs, followed by a computer-generated image, highlighting advanced synthetic image generation techniques. To tackle this issue, a field of research aimed at developing image manipulation detection and localization techniques has been designed to detect and locate manipulations in digital images automatically. Therefore, the field involves developing algorithms to analyze images for signs of tampering, including inconsistencies in lighting, texture, and other visual features, and identifying regions of an image that have been altered.

Among them, numerous approaches have been put out to solve the issue of unaware information forensics for source identification and authenticity of the content. Towards this, a significant subfield of image forensics is "image manipulation detection and localization," its goal is to pinpoint precisely where digital photo editing has taken place. Thus, it facilitates more accurate image analysis, which can shed light on the scope and nature of the alteration. Furthermore, methods for localizing image manipulation often compare image attributes, such as textures and edges, from different parts of the image to look for discrepancies that manipulation may affect. However, measures to increase image modification detection accuracy and efficiency are ongoing; hence, new techniques and tools are continually being created. Therefore, research in this field continues to detect and localize image manipulations.

In recent years, there has been a significant growth in the number of research works exploring data-driven approaches for image manipulation detection and localization. However, these methods utilize complex neural network architectures and large amounts of training data to learn and identify patterns in the images. Among them, some of the classical techniques [23, 31, 67, 154, 185], machine learning [54, 64, 118, 137], and deep learning [15, 37, 52, 198, 215] are used to achieve these goals. For better understanding, the generalized flow diagram for the image manipulation detection and localization approach is presented in Figure 2. Therefore, image manipulation detection and localization techniques provide a more comprehensive understanding of the nature and extent of image manipulation, which can be helpful in various contexts. To achieve this goal, this survey article will provide an exhaustive review of the recent developments in image manipulation detection and localization techniques. Additionally, it specifies the contemporary challenges and future tendencies of this field of research.

Table 1. Some of the Most Common Types of Image Manipulations

Image manipulation type	Description
Retouching	Image editing lead to changing its hue, saturation, and contrast to make it look better.
Cloning	It's common practice to cut off problematic areas of an image and paste them elsewhere.
Content swapping	Involves swapping out one component with another.
Addition of objects	Adding new objects to an image to create a false scene representation.
Removal of objects	Removing objects from an image to create a false representation of a scene.
Face manipulation	Altering facial features in an image, such as changing a person's expression or adding or removing wrinkles.
Background manipulation	Changing an image's background to create a false representation of the location or environment.
Compression artefacts	The degradation of image quality can occur during image compression.
Splicing	Combining parts of two or more images to create a new image that falsely represents the original scenes.

(a) (b) (c)

(d) (e) (f)

Fig. 1. Examples of different image manipulation attacks: (a) and (b) show source images used to generate a FaceSwap-manipulated image (c); (d) demonstrates a Deepfake-altered image; (e) and (f) are computer-generated synthetic images.

1.1 Motivation

Image tampering and processing detection systems were developed due to the increased occurrence of image alteration and the necessity for reliable techniques to assess digital image authenticity and integrity. Image tampering is a significant problem in forensics, law enforcement,

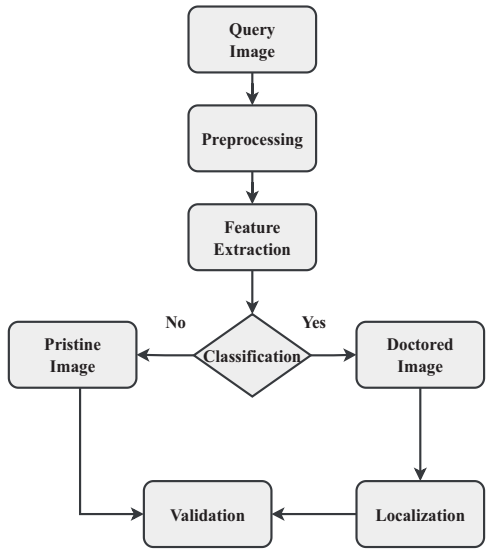


Fig. 2. Flowchart for generalized image manipulation detection and localization approach.

journalism, and multimedia content authentication because of the widespread availability of image-altering software and the ease with which digital images may be shared. Image alteration and processing detection methods are essential for validating a photo's authenticity in various domains, including forensics, law enforcement, and legal proceedings. Digital images must be reliable sources of evidence, and detection methods help assess whether an image has undergone manipulation. With the rise of online platforms and social media, images play a crucial role in news and information dissemination, necessitating sophisticated identification methods to prevent the spread of misleading or doctored visuals. Additionally, content authentication and copyright protection rely on image manipulation detection to identify plagiarism, copyright infringement, and unauthorized alterations, safeguarding intellectual property rights. However, while detecting manipulation confirms that an image has been altered, the absence of detected manipulation does not guarantee authenticity with absolute certainty. Some manipulations may go undetected due to technical limitations, emphasizing the continuous need for advancements in detection techniques. Motivated by this, this article includes a comprehensive review of detection and localization manipulations, including related surveys in this field. To contextualize this review, the following section explores contributions from prior surveys and research in the field of multimedia forensics.

1.2 Related Surveys

Numerous survey articles in the recently published literature have demonstrated their comprehensive perspectives on multimedia forensics or its subfields [14, 143, 155, 171, 174], providing a foundation for understanding the evolution of detection techniques. Towards this, initially, Farid [61] reviewed image manipulation detection systems, and divided them into five varieties: camera-based, physically-based, pixel-based, format-based, and geometric-based techniques. Furthermore, Qureshi et al. [149] and Birajdar and Mankar [24] gave a more thorough review of the advancement of image forgery detection methods. Along with image forgery detection, Rocha et al. [157] covered the subjects of source identification and steganalysis; later, similar reviews of image forensics were published [1, 14, 34, 143, 155]. Likewise, Stamm et al. [174] extended the scope of the forensic inquiry from manipulation detection and expanded to encompass the

identification of audio and video counterfeiting. However, a few survey publications specifically state that they focus on identifying image tampering when an area of the original and tampered images differ. In [76], passive image tampering detection methods are classified based on various levels of tampering indications, such as statistics of pixels and coefficients, features like sharp edges and lighting inconsistency, and semantic content. On the other hand, Christlein et al. [46] analyzed common copy-move detection methods with a more narrow focus. In a similar line, Zampoglou et al. [214] examined approaches for identifying and localizing splicing utilising multiple source images. Experiments were conducted by Schetinger et al. [162] to determine the effects of various splicing techniques on typical forensic traces.

Building on these studies, it is essential to recognize that image manipulations can occur at both pixel and content levels, each leaving distinct traces that can be analyzed for detection. Therefore, common picture alterations can be done at pixel and content levels, such as when an object is removed from a digital shot or when an image is resampled to make it smaller. Each alteration product has an original image, regardless of the level of procedures that were used. One can identify the modifications by comparing the predicted output to the original and distinguish the various subcategories of image manipulation based on the diverse patterns of these modifications. As a result, there are several studies on the topic of forgery detection; however, only a small number of studies have focused on deep learning-based manipulation detection techniques, including manipulation-specific such as resampling, median filtering, copy-move, splicing, and object removal, as summarized in Appendix A. However, there is a lack of comprehensive surveys that thoroughly examine all aspects of forgery detection and localization, including but not limited to definitions, classifications, models, performance evaluation methodologies, and ongoing research topics. Hence, a systematic review of forgery detection techniques based on deep learning is necessary, which covers the most recent developments in the forensic field.

1.3 Key Contributions

The present survey is based on the part and a small number of recent surveys [14, 19, 65, 221] that focus solely on manipulation identification in images. The survey then provides a comprehensive and well-organized analysis of data-driven approaches for image manipulation detection, with a key emphasis on the most recent state-of-the-art findings from the last decade. The following is an overview of the most important contributions of this study.

- A systematic and technical look at the current state-of-the-art image manipulation identification using data-driven techniques is provided, including both the performance and real-time processing approaches.
- To benefit researchers in selecting the most appropriate deep learning approach for a certain application, a comparison of existing methods has been provided.
- A comprehensive examination of performance evaluation approaches is provided, including details on datasets, computing infrastructure, different evaluation benchmarks, and performance measures for quantitative and qualitative analysis.
- In the qualitative analysis part, we provided a high-level overview of deep learning methods to reveal the underlying decision-making process.
- This review highlights research gaps and emerging trends in manipulation detection and localization, offering insights into future directions such as advanced detection methods, multi-modal analysis, and addressing evolving challenges in multimedia forensics.

The rest of the article is structured as follows. Section 2 described detection techniques for identifying manipulations and processing operations using classical and deep learning approaches. In Section 3, we go into the technical discussion of several manipulation-specific type fingerprints.

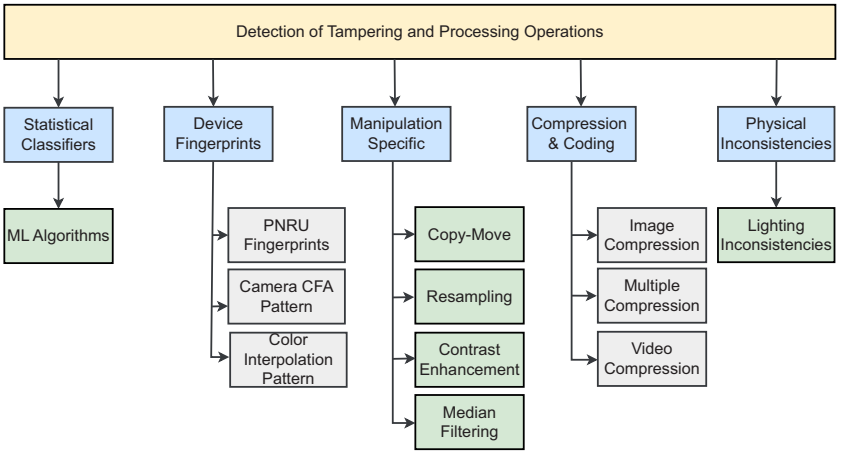


Fig. 3. The review of detection techniques for identifying image tampering and processing operations.

Furthermore, a technical discussion on localization techniques is reported in Section 4. Section 5 comprehensively investigates performance evaluation strategies, including a discussion of benchmark datasets, computational infrastructure, evaluation criteria, and performance measures for both quantitative and qualitative analysis. In Section 6, we highlight the difficulties researchers encountered when attempting to use various deep-learning techniques to detect forgeries in images. In Section 7, we draw some final findings and discuss some possible future uses.

2 Detection Techniques for Identifying Image Tampering and Processing Operations

Multimedia content from a suspicious or unidentified source often offers crucial details. For instance, a hostile foreign government may release photos or videos of an event that could have serious political or military consequences. Hence, we need more evidence before we can believe the events, such as any manipulation the multimedia material has suffered, must be identified, and the authenticity of the media must be established. To solve this, researchers have created a wide range of forensic approaches that function independently on extrinsic security measures due to the low probability that the multimedia content would have security benchmarks built into it before any processing takes place. Among these, some methods like softening the boundaries between foreground and background, adjusting the contrast, rotating, and resizing the images come under this category. A comprehensive overview of various detection techniques employed to identify image tampering and processing operations is illustrated in Figure 3. Consequently, many academic works centre on discovering these fundamental processes as an alternative for potential forgeries and reported two main types of image alteration detection methods, classical and data-driven approaches, which can be roughly represented by the two kinds of feature representation as shown in Figure 4 and 5: those that are manually created and automatically learnt.

2.1 Classical Methods for Manipulation Detection

In classical manipulation detection approaches, as illustrated in Figure 4, the pre-filtering step plays a critical role in forensic analysis by applying various filters to generate residuals that highlight manipulation traces. These residuals are texture-independent, making them particularly valuable for forensic applications. In this context, Popescu et al. [144] considered image resampling as a form of manipulation and proposed a method to identify traces left by resampling operations. In order to detect resampling, which is present whenever the image is rotated or resized, specific

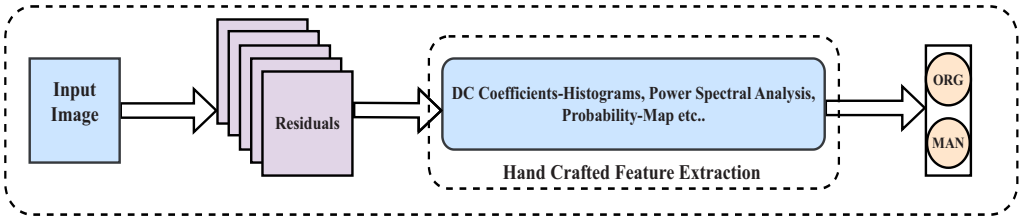


Fig. 4. Image manipulation detection using classical approaches has three phases: pre-filtering (noise residuals) followed by handcrafted feature extraction and classification.

approaches [95, 145] use periodic artefacts. Several methods look for inconsistencies in the blurring or anomalies at the edges of objects when a composition is being done [7, 59]. Image portions are frequently copied and moved to create duplicates or conceal details. Naturally, the existence of overlapping areas is a crucial indicator of forgery. Still, duplicates are commonly altered to suppress evidence, and near-identical natural items occur; hence, this makes forensic analysis more difficult. In [66], fundamental work on copy-move detection and much has been reported on this issue. Even with rotation, scaling, and other geometric distortions, copy-move detection can now identify effective and efficient algorithms [46]. Keypoint-based methods are highly effective [4, 167], while dense-field approaches [50, 158, 211] are more precise and can handle occlusive attacks. Dense-field techniques have also been demonstrated to be successful in detecting inpainting [48].

The principles of machine learning construct these techniques, such as a classifier being trained on many examples of unaltered and altered images, once suitable characteristics are identified that aid in discriminating between the two. It is important to note that the forensic analyst carefully crafts features based on a thorough familiarity with the intended changes. Specific artefacts are generated by double JPEG compression [76, 88] or those artefacts that are related to the **camera response function** (**CRF**) [81, 117], can be identified using features that have been developed for this purpose. However, these universal features can detect multiple forms of tampering and are based on appropriate image statistics. Selecting features with the most discriminatory power might be challenging, but strong statistical models for natural photos can help. Removing the semantic image information, which can be ignored as noise, is the first step towards highlighting statistical anomalies produced by modifications, which has already been noticed in [18]. As a result, the multiple valuable characteristics are often recovered from noise residuals [109, 219]. There is a lot of room for improvement in features derived from high-order image statistics, as demonstrated by the groundbreaking work of Farid and Lyu [62]. These characteristics excel in many application domains, including computer graphics, biometrics, and steganalysis, since they detect minute changes in an image's micro-textures. Furthermore, the most well-known rich models [67] for identifying noise residuals were first developed for steganography and have since been successfully implemented in forensics. After passing the image through a series of high-pass filters, which can pick up slightly different artefacts, thus, these features are then constructed based on how often specific neighbours are related.

2.2 Deep Learning Methods for Manipulation Detection

On the other hand, in deep learning-based methods, as depicted in Figure 5, feature extraction is performed automatically, eliminating the need for manual pre-filtering steps and generating feature maps. Therefore, these models are the finest approaches for feature learning and classification-type research problems [79]. Over the past decade, data-driven techniques have been applied broadly across disciplines, and the adoption rate of data-driven models has accelerated.

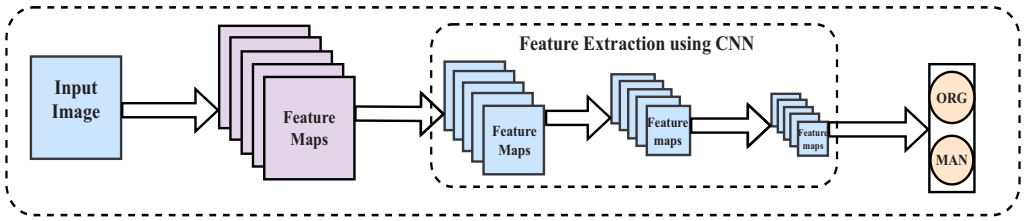


Fig. 5. Image manipulation detection using a deep learning-based approach has three phases: pre-filtering (feature maps) followed by automatic feature extraction and classification.

Images manipulated are often detected with the help of data-driven methods [15, 37, 151, 159]. For tamper detection, a large dataset of authentic and altered images is fed into deep learning-based image manipulation detection models. In a short amount of time, the model can tell the difference between two categories when a manipulated image is inputted to it [37]. Both training and testing photo sets construct a dataset, and the underlying features of photos are gathered with the help of a competent training model. Furthermore, a data-driven trained model employs numerous layers, each with its trainable parameters, to classify photos. However, more data for training a data-driven model yields more accurate results, yet training with a huge volume of data is labour-intensive. Therefore, one possible solution involving using computer-generated visuals in model training will reduce the complexity reported in [102, 168].

On the contrary, in the early days of trying to spot photo manipulation was only sometimes taken into account that one of the three tampering techniques (i.e., copy-move, copy-paste, and inpainting) would be followed by post-processing of the image. For illustration, copy-paste was described as the straightforward procedure of copying and pasting image portions from one image onto itself or another without employing post-processing [76]. To make the finished artefact less visually suspectable, realistic tampering frequently incorporates post-processing techniques to smooth the boundaries of tampered regions. Post-processing processes come in two different varieties. Firstly, active post-processing techniques enhance the tampering effect, such as blurring, resampling, adjusting brightness, and adjusting contrast in an image. Secondly, passive post-processing, which could accidentally be applied to altered photos when sending data compressing, adding noise too, and de-saturating images using JPEG [181].

3 Technical Discussion on Manipulation Detection Techniques

Automatic manipulation detection of doctored images and videos is challenging because of the complexity involved during manipulation creation. In this section, a variety of manipulation detection methods, including detection and manipulation parameter classification, are discussed. In addition, manipulation-specific forgeries such as resampling, **contrast enhancement** (**CE**), median filtering, compression and coding fingerprints to detect single, double, and triple compression and a combination of these are discussed.

3.1 Resampling Fingerprints

Initially, we started with resampling manipulation, which occurs whenever an image is enlarged, rotated, or subjected to an affine transformation; it is of particular forensic significance. However, finding resampling artefacts throughout an image indicates image processing, but it does not prove that the image is a forgery. Generally, scaling or rotating the object added to the backdrop image is typically required to make the fake appear aesthetically realistic when employing a cut-and-paste technique. For instance, any combination of non-aligned cropping, scaling, rotation, and sequential

geometric modifications can be performed on a digital image I. According to sampling theory, designing a *sinc* filter with infinite support is tough to yield perfect reconstruction. Therefore, many other interpolation filters with finite support, such as linear, cubic, and truncated *sinc*, are widely used. Hence, an affine transformation of the axes $(x_1, x_2) \in Z^2$ to $A(x_1, x_2)^T + (\theta_1, \theta_2)^T$ will also produce a new sample grid with different pixel intensity values. The 2×2 transformation matrix (A) will be used to expand, rotate, and skew the grid, while the translation matrix $(\theta_1, \theta_2)^T$ will be used to move the grid. The transformation matrix for the scaling process looks as follows

$$A_\zeta = \begin{bmatrix} \zeta & 0 \\ 0 & \zeta \end{bmatrix}, \tag{1}$$

where, 'ζ' is a scaling factor. For image rotation, the A_θ with θ rotation angle is given as

$$A_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{2}$$

when resampling takes place, the new sampling grid of the image (y_1, y_2) and its intensity values can be obtained by performing convolution with interpolation kernel $h(x_1, x_2)$

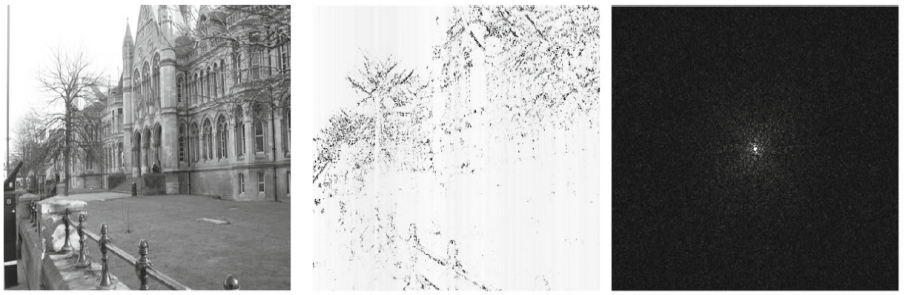
$$I_R = h(x_1, x_2) * I(x_1, x_2), \tag{3}$$

where, I_R is a doctored image created from the original image I. Therefore, it is essential to consider resampling and post-JPEG compression combinations, as they can provide crucial information about the image's prior history.

Toward this goal, several authors proposed resampling detection algorithms to detect forgeries in images. In [145], the authors discussed that every pixel in a resampling image is correlated with its neighbouring pixels. They used an **Expectation/Maximization** (**EM**) algorithm to catch the periodic relationship among the image pixels. This algorithm is based on a series of baseline variables and takes a long time to get over. Kirchner [95], presented a resampling detector based on the highest spectral slope probability map *(p-map)* in a linear predictor to reduce computational complexity. However, this detector got better results in upscaling but could not detect downscaled images. Gallagher [68], Mahdian and Saic [128] have proved that second-order derivatives in the interpolated signal exhibit periodicity. This feature is computed by estimating the DFT of an averaged signal. However, the major limitation of this feature is that the performance is decreased in a compressed scenario.

Kirchner et al. [97] presented a method to detect the resampling of compressed images using JPEG pre-compression artefacts. The techniques used for detecting and estimating the resampling factor of JPEG were refined by combining the energy-based [64] and predictor-based methods [142]. Later, Bianchi and Piva [23] estimated the quantization matrix and resampling factor in recompressed images using the reverse engineering strategy. Moving from the spectrum-based methods, Padin et al. [184, 185] adopted other linear and 1-D signal methods. Furthermore, in [147], presented a resampling detector that can detect the difference between upscaled and genuine images by utilizing a periodicity of probability map (P_{map}) frequency spectrum as shown in Figure 6. However, this resampling detector is analyzed using the upsampled images while leaving downscaled analysis for later use.

On the other hand, CNNs are widely used in media forensics due to their development in computer vision [77, 101, 168]. Unlike traditional vision problems, most image forensics research focuses on low-level patterns that are more problematic than anything else. This philosophy is commonly used in deep learning setups. Many current deep-learning approaches start by feeding residual images into a pool of networks, constraining the learning process to learn a residual layer. This objective can be attained by adding an initial layer of fixed weights at the first layer [152].



(a) Original image, P_{map} and its periodic Fourier spectrum

(b) Resampled image ($\zeta = 1.5$), P_{map} and its periodic Fourier spectrum

Fig. 6. Detection of resampling by utilizing probability map *(p-map)* as reported in [147].

Furthermore, trainable [123] or highpass filters [15] are utilized for general-purpose image manipulation detection. The main drawback of their approach is that they could only analyze upscaling scenarios with low JPEG compression, and downscaling with high compression was left for future work.

In [197], 385 manipulations are learned for feature extraction and to localize the image tampering as an anomaly detection problem. Furthermore, Xception-based CNN architecture is implemented to classify multiple image manipulations by focusing on small-size image patches [204]. Furthermore, MCNet exploits spatial, frequency, and compression multi-domain features to classify JPEG images [207]. However, these networks achieve satisfactory results for pre-JPEG compressed images, but the results are significantly degraded for JPEG post-processing (i.e., recompressed images). MSRD has recently been utilized to learn the features of multiple image manipulation detection [150]. Furthermore, for better understanding, the experimental settings such as interpolation techniques, feature extraction techniques, and classifiers of the reviewed works on resampling detection and estimation using both classical and deep learning approaches are summarized in Table 2.

3.2 Median Filtering Fingerprints

The forensics community also pays a lot of attention to median filtering, another editing procedure. Median filtering is nonlinear; therefore, it can smooth a signal without distorting its edges. For this reason, median filtering is frequently employed to smooth or denoise digital photographs. The median filter, initially described by Tukey in [86], is by far the most often used nonlinear filter. Since then, a median filter has been used in other image-processing contexts. For the detection of median filtering, the statistical analysis of median-filtered images in two dimensions is analyzed

Table 2. Summary of Reported Techniques for Image Resampling Detection by Conventional and Deep Learning Methods

Literature	# Images	Interpolation technique	Feature extraction technique	Classifier
Popescu and Farid [145]	200	Bicubic	Probability maps (*p-map*) spectrum	Expectation/Maximization (EM) Algorithm
Mahdian and Saic [128]	40	Bicubic	Variance of n^{th} order derivative of image using Radon Transform	Threshold-based peak detector
Kirchner [95]	200	Bilinear	Probability maps (*p-map*) spectrum	Cumulative periodograms
Feng et al. [64]	7,500	Bicubic	Normalized energy density (DFT)	SVM classifier
Bianchi and Piva [23]	500	Bilinear	Integer periodicity property (IPM)	Threshold-based peak detector
Padin et al. [185]	1,317	Bilinear, Bicubic Lanczos	Asymptotic eigenvalue distribution of sample autocorrelation matrix	Threshold-based peak detector
Qiao et al. [146]	1,338	Bilinear and Nearest neighbor	Pixel level texture map, Block level texture map, Region level texture map	LRT test
Bayar and Stamm [17]	3,334	Bilinear	Constrained CNN	ET Classifier, FCN
Cao et al. [32]	50,000	Bilinear	Horizontal, vertical and interleaved stream	FCN
Liang et al. [112]	1,700	Bicubic	Residual neural network	FCN
Yang et al. [203]	13,800	Bilinear	Magnified layer and CNN	FCN
Liu et al. [120]	1,445	Bilinear	VGG Net with 25 convolutional layers	FCN
Yu et al. [207]	80,005	Bilinear,Nearest neighbor, Bicubic	Visual artifact network, Compression artifact network	Ensemble Classifier
Rana et al. [150]	5,500	Bilinear	Multi-scale residual module, feature extraction blocks	FCN

in [114]. As part of a pixel-based image operation, the median filter processes each individual pixel in the image separately, and the median filtered image I_{med} is created as follows:

$$I_{med} = med_n(I(x, y)), \tag{4}$$

where, med_n in Equation (4) is median filter with window size $n \times n$, and the standard filter window size is 3×3, and 5×5 pixels. Liao et al. in [114] is the first to investigate the connection between the input and output of an MF for a 2D signal, and it may be of welfare from a forensics perspective. In [29], Alan Conrad Bovik explored how a median filter altered the variance in pixel brightness. The study was conducted using iid signals, which have no connection between neighbouring pixels, despite the fact that in real-world digital images, there is a high association between neighbouring pixels. Although image processing primitives may not always compromise an image's authenticity, they are still of interest during a forensic analysis of photos because of the different ways in which they might alter the methodologies used to analyse the image. Forensic scientists are particularly interested in basic image processing techniques like the **median filter** (**MF**). In the context of forgeries, MF is useful for hiding the obvious signs of manipulation left by other processes. Many forensic techniques assume a linear connection between neighbouring pixels, such as resampling and CFA interpolation detection [145]. However, using a local linear predictor of pixel intensities, Popescu's resampling detection method (in [145]) has been demonstrated to be susceptible to median filtering (in [96]). As a nonlinear operator, an MF can be used

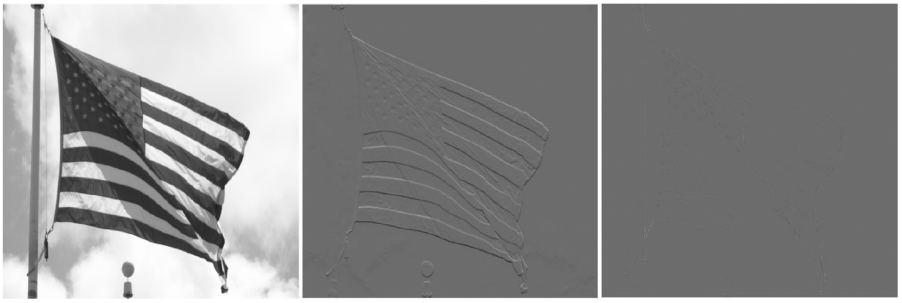


Fig. 7. Example showing its first order and its MFR [93].

as a weapon against interpolation detection techniques by destroying such linear dependencies. Therefore, new resampling activities can be uncovered using forensic MF detection, such as the extraction of the **median filter residual** (**MFR**).

$$MFR = med_n(I) - (I). \tag{5}$$

Most of the techniques, such as conventional methods and CNN-based methods, used MFR from the test image [74, 93, 124, 170, 187]. For illustration, in [93] discussed the difference between first order and MFR as shown in Figure 7. The final step in forgery is to use a median filter to mask any signifying edges. In addition, many modern forensic techniques assume a linear relation between pixels. Still, introducing a nonlinear filter like the median filter eliminates this linearity. Median filter forensics detection techniques can be roughly categorised according to the domain from which image features are derived. Median filtered image detection was accomplished using a modified CNN by Chan et al. [37]. Furthermore, the authors argue that adding the new filter (i.e., MFR) layer enhances the effectiveness of the standard CNN. In [15], Bayer et al. present a broad strategy for identifying simple operator manipulation. After training a CNN to suppress image contents, the trained CNN automatically extracts features from the images using the prediction error filters learned by a new constrained convolutional layer. Tang et al. [180] constructed a novel MFNet network to detect MF forensics in low-resolution photos. Furthermore, in their proposal for a CNN-based technique for recognising low-resolution MF images in post-JPEG compression, Yu et al. [210] suggested using 20 **High pass filter** (**HPF**) residues to conceal image contents during the pre-processing step. Similar to [210], Yang [204] implemented a CNN with a magnified layer as part of the pre-processing phase to detect MF. Recently, an adaptive filtering layer in the discrete cosine transform domain has been presented as a forensics technique to detect MF by Zhang et al. in [216]. Finally, the total number of images for experimentation, image size, the feature extraction techniques, and types of ML classifiers employed in some of the most influential publications on median filter forensics are summarized in Table 3.

3.3 CE Fingerprints

The lighting of images can be adjusted with the help of an editing method called CE. The forger may boost contrast in manipulated parts of an inspected image to guarantee consistent illumination throughout a cut-and-paste forgery. To change contrast, monotonically increased and non-linear mapping is applied to the signal values. In general, the pixel values of the test image (I) inherently changed using the mapping function ($T(.)$)

$$y = T(x),\ x, y = 0, 1, 2, 3, \ldots\ldots 255, \tag{6}$$

Table 3. Summary of Reported Techniques for Image Median Filtering Detection by Conventional and Deep Learning Methods

Literature	# Images	Patch size	Feature extraction technique	Classifier
Yuan et al. [212]	5 databases, 10000	64 × 64, 32 × 32 16 × 16	DBM, OBC, QGL, DBC, FBC Feature set, Fusion feature set	C-SVM Classifier
Kang et al. [93]	6690	128 × 128, 64 × 64 32 × 32	MFR, Auto regression model	C-SVM Classifier
Zhang et al. [218]	1000	256 × 256, 128 × 128 64 × 64	High-order local ternary patterns, Kernel PCA	C-SVM Classifier
Chen et al. [37]	15352	64 × 64, 32 × 32	MFR-based CNN, Filter layer	FCN
Liu et al. [118]	4014, 1338	128 × 128, 64 × 64	Frequency domain Feature, Annular accumulated points (AAP)	C-SVM Classifier
Niu et al. [137]	1338	256 × 256, 128 × 128 64 × 64, 32 × 32	Rotation invariant uniform local binary pattern (LBP), Pixel Difference Matrix (PDM)	C-SVM Classifier
Tang et al. [180]	14800	64 × 64, 32 × 32	Magnified layer, Mlpconv layers	FCN
Wang et al. [186]	933, 1338, 10000	64 × 64, 32 × 32 16 × 16	DCT subband energy distribution feature	C-SVM Classifier
Luo et al. [124]	30000	32 × 32	MFR, Filter layer, ResNet	FCN
Guptha et al. [74]	5352, 1309, 13985	384 × 512, 256 × 256 128 × 128	MFR, Pearson parameter (ratio of Skewness and kurtosis)	C-SVM Classifier
Shan et al. [164]	10000	64 × 64, 32 × 32	Deblocking layer, Fused Filtered Residual (FFR)	FCN
Zhang et al. [216]	30000	128 × 128, 64 × 64 32 × 32	DCT-CNN, Adaptive filtering layer, Multi-scale feature extraction	FCN
Wang et al. [187]	27000	256 × 256	MFR, Quantunion CNN	FCN

where x, y represent pixel values before and after applying the mapping function. To determine whether a digital image has been subjected to CE processing, many methods [31, 54, 173] have been proposed. Earlier CE forensic techniques [173] compared 1D grayscale histograms of the original and edited images and found discrepancies between them. According to Stamm and Liu [173], a pure image's 1D histogram has a smooth envelope, while a contrast-enhanced image's envelope is ragged and contains peaks and troughs. The histogram is Fourier transformed, and its power spectral density is determined. When is greater than a predetermined threshold, we know that the target image has been contrast boosted. An alternative interpretation of the CE image's 1D histogram was presented by Cao et al. [31], is evaluated the zero-height gap in its histogram as the imprint of CE because different image processing approaches like JPEG compression might cause histogram peaks. The 1D histogram is examined to determine the number of gaps between two bins, and if the gap bins are greater than a predetermined threshold, the image is subjected to CE. However, processing with the counter-forensic approaches makes it difficult to erase these elementary 1D histogram-based features.

In order to solve this issue, De Rosa et al. [54] attempted to categorise the photos as either being intact, contrast-enhanced, or anti-forensically attacked by presenting an SVM. For illustration, the authors use a **general linear combination model** (**GLCM**) of the original, gamma-corrected, and CE images, which are depicted in Figure 8. In this approach, classification is performed using a three-class SVM, with a one-dimensional histogram created from summing variance values of each row of GLCM serving as a feature vector for SVM training. Since the GLCM was utilized to construct a separate elementary feature vector, it is hard to say that the 2D data was used for acceptable purposes. It's an uncertain condition; this is the first time second-order statistics have been applied

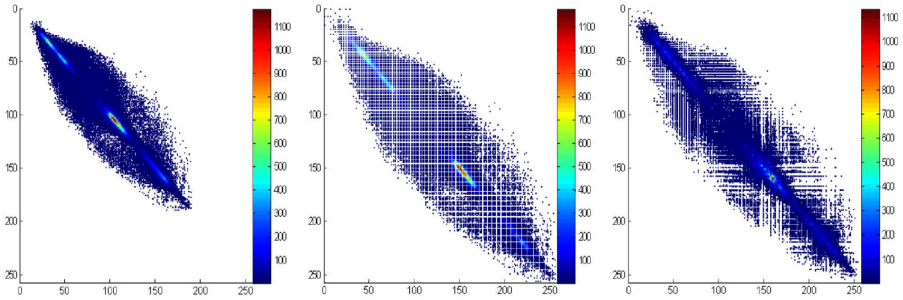


Fig. 8. GLCM of the original image, gamma corrected, and enhanced image after remapping attack presented in [11] is applied by the approach in [54].

Table 4. Summary of Reported Techniques for Image CE Detection by Conventional and Deep Learning Methods

Literature	# Images	Patch size	Manipulation type					Feature extraction technique	Classifier
			GC	HS	CLAHE	HE	SM		
Stamm et al. [172]	341	Various sizes	✓	✓	-	-	-	Image histogram and frequency domain	Threshold test
Cao et al. [31]	3 databases	Various sizes	✓	✓		✓	-	Zero-height gap, Peak/gap similarity measure	Threshold test
Lin et al. [116]	100	1600×1200	✓	-	-	✓	-	Inter channel similarity matrix	Threshold test
Alessian et al. [54]	500	512×384	✓	✓	-	-	-	GLCM, Second order statistics	C-SVM Classifier
Hareesh et al. [154]	5,000	512×512	✓	✓	-	-	-	Anti-forensic approach	C-SVM Classifier
Sun et al. [175]	5,000	256×256	✓	✓	-	-	✓	GLCM input layer, CNN	FCM
Barni et al. [10]	2×10^6	64×64	✓	✓	✓	✓	✓	CNN	FCN
Zhang et al. [215]	16,000	256×256	✓	-	-	✓	-	Histogram as input and 1D CNN	FCN
Shan et al. [164]	10,000	256×256	✓	✓	-	-	-	GLCM layer, cropping layer	FCN
Zou et al. [227]	80,000	128×128	✓	-	-	-	-	GAN, GLCM domain and histogram domain	FCN

to CE forensics, but it certainly is the first time it has been attempted here. The previously reported CE forensic approaches [31, 54, 173] are based on handcrafted features and include different feature extraction and classification. As a result, they produce unsatisfactory results in detecting images that have been modified in counter-forensic attacks. While analysis based on deep learning models for CE forensics is lacking, several attempts have been made to handle various alterations in digital image forensics by employing deep learning models. CNNs were initially used in [37], one of the earliest publications on image forensics. The authors of this research created a median-filtered image detector using a CNN. In this scenario, a pre-processing step was performed to obtain the MFR and better show the manipulation traces. Providing this residual to the CNN allowed the detector to enhance the manipulation-related characteristics, leading to improved performance compared to training with an image as the input directly. As an alternative to the traditional pre-processing stage, Bayer and Stamm [15] developed a constrained pre-processing layer while simultaneously concealing the image information, preventing the learning of content-dependent characteristics. However, it is not just median filtering that this technique can handle; it can also handle filter-based modifications like bilinear interpolation, Gaussian blurring, and the addition of AWGN. In addition, Table 4 summarises the experimental settings, including interpolation methods, feature extraction strategies, and classifiers, of the examined studies on CE detection and its estimation for easier comprehension, which employ conventional and deep learning approaches.

3.4 Compression Fingerprints

Compression and encoding leave their recognisable fingerprints, just as many other signal processing procedures. Therefore, it's common to find compression or coding fingerprints in digital

multimedia signals. Due to their prevalence, researchers have devised numerous methods for compressing or encoding fingerprints to carry out various forensic functions. Some examples include checking the legitimacy of a multimedia item, finding out its origin, and processing history. In the field of forensics, lossy compression clues have received the greatest attention and research. Due to the overall use of JPEG as a compression standard, its fingerprints play a crucial role in forensics. It is an important and well-studied forensic problem to specify whether an image has been compressed using JPEG twice with differing quality parameters or not. For instance, proof of double JPEG compression indicates that an image may have been manipulated, but after that, it must be re-saved after being modified and reveal essential details about its processing history. Visual and noise elements work together to disclose tampering artefacts. Double JPEG compression is infamous for the anomalous artefacts it leaves behind in the DCT domain, most notably on block-DCT coefficient histograms [144]. For this reason, many of the reported detection techniques rely on the first-order statistics of DCT coefficients to make their decisions. As an example, the model-based approaches in [100, 177] rely on the distribution of the first or second-order effective coefficients in the block-DCT, and the data-driven approach in [141] analyses the histograms of the coefficients at low frequencies. Detectors that are data-driven and use features obtained from second-order statistics have also been presented (see, for example, [36, 188]). While many of these methods are effective at determining whether an entire image was compressed using single or double JPEG. For this, they often need to catch up with smaller blocks because of the problem of accurately calculating the statistics involved. Since only a tiny portion of the image has been altered, they are ineffective in a tampering detection scenario.

Many alternative techniques for identifying non-aligned double compression (NA-DJPEG) have been presented, with some of them drawing on multiple features retrieved from the pixel domain [43, 125] and others from the DCT domain [21, 148]. For instance, a method for detecting aligned and non-aligned re-compression was proposed by the authors of [43]. The plan effectively combines periodic artefacts in spatial and frequency domains. In particular, when NA-DJPEG compression is present, a collection of features is computed to measure the periodicity of blocking artefacts (BAR), and when aligned double compression (A-DJPEG) is present, a different collection of features is utilised to estimate the periodicity of DCT coefficients. Compared to [21], this method is inferior for detecting non-aligned re-compression. Also, Bianchi and Piva [22] provide a forensic technique for localization tampering in the context of DJPEG compression, whether aligned or not. Wang et al. [190] proposed a CNN-based double JPEG compression detector with a histogram consists 99 bins as input. The DCT coefficients were then organised in a zigzag pattern across the first nine AC sub-bands, creating nine corresponding histograms. Only coefficient values between -5 to +5, including "0", are considered when accumulating to a histogram for each sub-band. In addition, to understand the histogram analysis, single, double compressed image histograms are depicted in Figure 9. Therefore, the histogram with 99 bins was obtained by concatenating the nine histograms, each with 11 bins. In [5], a multi-stream CNN that can distinguish between uncompressed, single-compressed, and double-compressed images is introduced. Furthermore, the multi-streams, with spatial-stream CNN, take a colour image with three channels, and the frequency-stream CNN is in a one-dimensional histogram of the DCT coefficients. Finally, the decision on JPEG compression is based on combining two feature vectors derived from the two CNNs. However, these two strategies employ 1D CNN; its input is not an image, and it is trained using the DCT coefficient histogram. Furthermore, using DCT coefficients forms images as input, Barni et al. [9] suggested a CNN-based non-aligned and aligned JPEG compression detector. In this, three separate inputs are fed as pre-processing steps (i.e., the mean subtracted image, the noise residual image, and the DCT feature image) and were tested to determine their effect on double JPEG lossy compressed images. Huang et al. [84] reported

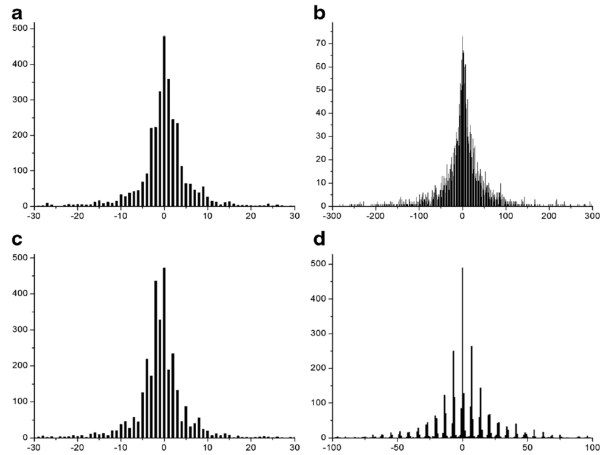


Fig. 9. DCT coefficient histograms at (0,1) position for single-compressed images with a) $QF1 = 60$ and b) $QF1 = 90$. Histogram of double-compressed image with c) $QF1 = 90$, $QF2 = 60$, and d) $QF1 = 60$, $QF2 = 90$ carried from [190].

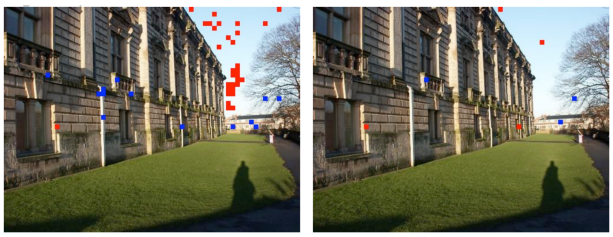


Fig. 10. Detected rounding (in red) and truncation error (in blue) blocks for (a) single and (b) double JPEG compression with QF = 90 in [84].

rounding and truncating errors as input to dense CNN, which can be able to identify single and double-compressed images with the same quantization matrix shown in Figure 10. Furthermore, a summary of the reported techniques developed for detecting double JPEG compression is provided in Table 5, highlighting the various methodological advancements and their performance characteristics.

3.5 Universal Manipulation Detection Techniques

Typically, several image tampering operations are detected using a universal image modification detection approach. Detecting the traces left behind during post-processing procedures is usually the foundation of universal image manipulation detection systems. By utilizing these traces, it is possible to identify universal image alteration using several deep learning-based techniques. Therefore, brief descriptions of the various universal image alteration detection methods are provided in Table **??**. It also includes the feature extraction, publication year, methodology, article summary, performance metrics, and evaluated dataset. In this section, we look back at recent related work for detecting and pinpointing tampering using CNNs. From a forensics perspective, the local reliance of pixels with their neighbours is the most crucial detail extracted by this pre-processing layer. Towards this goal, a general solution for detecting image alteration using a deep learning approach is provided by Bayar and Stamm [15]. Specifically, in [15], the k filter weights individually specified

Table 5. Summary of Reported Techniques for Double JPEG Compression Detection Methods

Literature	# Images	Patch size	Quality factor		Misallianment		Feature extraction technique	Classifier
			$QF_1 = QF_2$	$QF_1 \neq QF_2$	Aligned	Non-Aligned		
Li et al. [104]	40,000	256 × 256	-	✓	✓	-	Multi-branch CNN, Histogram features, Slice layer	FCN
Barni et al. [9]	3,840K	256 × 256	✓	✓	✓	✓	DCT histogram, CNN	FCN
Huang et al. [84]	11,338	512 × 512	✓	-	✓	-	Traditional handcrafted features, Dense CNN, Fusion	SVM Classifier
Wang et al. [190]	132,000	Various	✓	✓	✓	-	1D CNN, DCT AC coefficients	FCN
Amerini et al. [5]	57,792	64 × 64	-	✓	✓	-	Multi-domain CNN (Spatial domain + Frequency domain)	FCN
Taimari et al. [178]	1,096	256 × 256	✓	✓	✓	-	Benford-based features, PCA	SVM Classifier
Deng et al. [55]		Various	✓	✓	✓	-	Multi-scale module, dicriminative module	FCN
Li et al. [105]	40,000	Various		✓	✓	-	Dual stream (DCT + spatial domain)	FCN
Niu et al. [135]	3,466	Various	✓	-	✓	-	Backward quantization error, statistical features, error based features	SVM Classifier
Park et al. [139]	1,140,430	Various	-	✓	✓	-	DCT histogram feature +2D CNN	FCN
Hussain et al. [85]	51,300	Various	-	✓	✓	✓	DCT layer, CNN	FCN

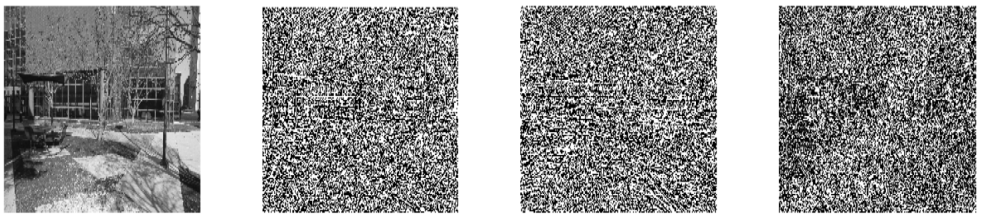


Fig. 11. Output of three filters utilized in constrain convolutional layer [15, 17].

as

$$\omega_k^1(0,0) = -1 \; and \sum_{x_1,x_2 \neq 0} \omega_k^1(x_1,x_2) = 1, \tag{7}$$

where, $\omega_k^1(x_1,x_2)$ denotes the weight at position (x_1,x_2) of the k th filter and $\omega_k^1(0,0)$ indicates the weight at the middle of the corresponding filter and its output is specified in Figure 11. The process is repeated for each pixel in the patch by shifting the kernels over the image patch. This new convolutional layer is an improvement over the standard CNN layers. For several image manipulation operations, general-purpose detection techniques have been developed [15, 17, 150, 207]. In contrast to previous methods, which needed manual preprocessing or feature selection, this method uses a novel convolutional layer to automatically hide the image's content while simultaneously capturing the traces left behind by the tampering operation. The layer's pixel's relationship to its immediate neighbours is that one being altered to hide the image's content. To learn the prediction error instead of image content by utilizing the convolutional layer, its initial weights are determined randomly, and the constraint is applied uniformly across all filters and iterations.

These methods have proven that they can autonomously learn image editing features from data. In order to identify the numerous image manipulation operations while hiding the image texture information, [15] proposes a unique restricted convolutional layer-based CNN, and [17] refines and improves this network. In addition, a densely connected CNN is reported for use in general-purpose visual forensics [41]. As part of this, high-pass filtering with the isotropic

convolutional layer makes the artefacts of image processing procedures stand out. Furthermore, in [131], a method for detecting image modification is provided based on [15] and employs a deep siamese CNN network. Instead of detection, they focused on determining whether or not a given set of input patches (two photos) had been similarly processed. Furthermore, they classified whether the image was manipulated or not. On the other hand, in [204], the Xception architecture is used for classifying numerous image processing tasks while considering small-sized images. The primary focus of this magnified image is to learn the altered characteristics of the modified image. Another work by Bayar and Stamm [16] provides a data-driven strategy for estimating manipulation parameters without requiring separate analyses of each form of manipulation. In [30], two techniques are employed to spot and pinpoint instances of photo tampering. In [8], Radon, Laplacian, and **Fast Fourier Transform** (**FFT**) characteristics train a deep neural network to identify altered images in one stream. The second stream employs a **Long short-term memory** (**LSTM**) network to learn the correlation (i.e., border modification between neighbouring blocks and the existing block of scaling characteristics), thereby supplying the discriminative characteristics to a softmax activation for classification. Contrarily, [217] uses a stacked autoencoder for feature understanding and employs contextual data to identify numerous image-altering attempts.

Moreover, effective data-driven methods have been developed for specialised image forensic tasks rather than including a pre-processing layer. The authors of [28] adopted downsampling with low-quality JPEG compression techniques to identify global modifications made to an image. In addition, denoising (content-sensitive low-pass filtering), low-pass filtering (blurring), high-pass filtering (sharpening), and tone correction (including histogram equalisation) were identified as the order of image processing history. In [27], the CNN-based architecture with the maximum-likelihood detector in terms of forensic accuracy is discussed. On the other hand, a dense CNN structure to improve the propagation of features relevant to image modifications for general forensics purposes is discussed [40]. Furthermore, it is manually designed, simultaneously detecting 11 distinct types of image modifications. In addition, some of the most important image manipulations ('MF': Median filtering, 'GB': Gaussian blur, 'RS': Resampling, 'JPEG': Lossy compression) and operator chain detection utilising data-driven algorithms are summarized in Table 6.

3.6 Technical Discussion on Manipulation Localization Techniques

This section discusses early reported manipulation localization methods and their performance on different datasets. Since different researchers utilise various datasets, imaging modalities, different segmentations, and validation criteria, comparing the performance of all these approaches is a difficult task. Forensic methods used to reconstruct the steps used to create a multimedia product, look for signs of tampering, or spot fakes fall into one of three broad types delineated by their respective modes of operation. The generalized forgeries and their corresponding ground truths are shown in Figure 12 such as splicing, copy-move, and object removal (inpainting) taken from DEFACTO dataset [129].

A forger may use a technique to change an image by swapping out one section for another section of the identical image. This commonly conceals an object's presence by masking it with naturally occurring textural elements like trees or grass. Furthermore, it can also make multiple copies of a critical feature in the image. Therefore, copy-move forgeries were one of the first forms of fraud studied by forensic experts. The scenario in which a region is replicated within an image is referred to as a "copy-move" in the technical literature. On the other hand, splicing, as opposed to the word "cut-paste", is used to describe region duplication between two or more photographs in published studies [78, 151, 163]. No matter how many photos are utilized as sources, the terms "splicing" and "composition" have been used to describe the two types of manipulation achieved by

Table 6. Summary of the Literature for Detecting Multiple Manipulations and Operation Chains Using Deep Learning-based Approaches

Year	Literature	Manipulation type				Feature extraction technique	Classifier
		MF	GB	RS	JPEG		
2016	Chu et al. [47]	-	✓	✓	✓	A mutual information-based criterion PSNR Noise Energy Pattern	Peak detector
2017	Chen et al. [44]	-	-	✓	-	Transformed block artifacts (TBAG), DCTR	SVM classifier
2018	Bayar and Stamm [17]	✓	✓	✓	✓	Constrained CNN	ET classifier, FCN
2018	Boroumand et al. [28]	✓	✓		✓	VGG Net	FCN
2020	Chen et al. [42]	-	✓	✓	✓	Feature extraction module, Transitional module	FCN
2020	Liao et al. [115]	✓	✓	✓	✓	Dual-Stream VGG Net	FCN
2020	Yu et al. [207]	✓	✓	✓	✓	Visual artifact network, Compression artifact network	Ensemble classifier
2021	You et al. [206]	✓	✓	✓	✓	Hierarchical Feature Extraction and Source Language Construction and Embedding	Transformer Encoder /Decoder
2022	Chen et al. [45]	✓	✓	✓	✓	Attack angle, mutual information scale	SVM classifier
2022	Rana et al. [150]	✓	✓	✓	✓	Multi-scale residual module, feature extraction blocks	FCN
2022	Li et al. [111]	✓	✓	✓	✓	Hierarchical Feature Extraction, Tokenization and Source Sentence Construction	Transformer Encoder /Decoder
2022	Kadha et al. [92]	✓	✓	✓	✓	Res-Net	FCN
2023	Chen et al. [38]	✓	✓	✓	✓	Feature coupling, Feature decoupling algorithm	SVM classifier
2023	Kadha et al. [91]	✓	✓	✓	✓	Res-Net, Compression feature extractor	FCN
2024	Samanta et al. [160]	✓	✓	✓	✓	SmartHash Algorithm	FCN
2024	Niu et al. [136]	✓	✓	✓	✓	Spatial Artifact Stream, Noise Residual Stream	FCN
2024	Singh et al. [169]	✓	✓	✓	✓	Constrained noise-view, Multi-scale feature learning	FCN

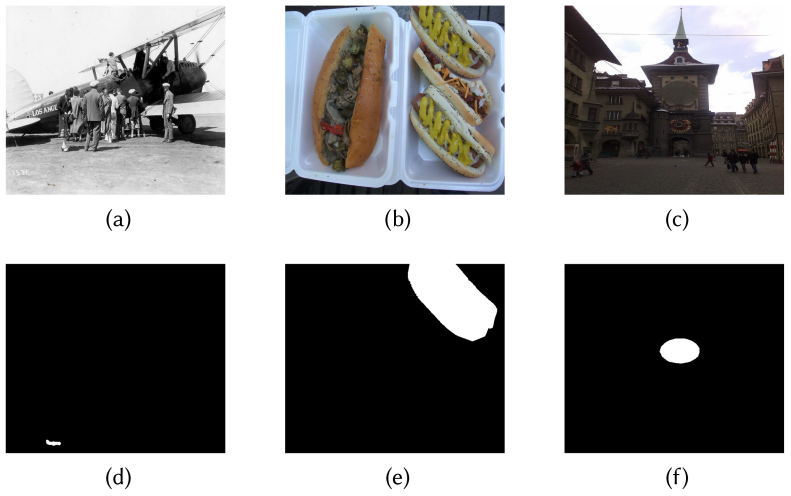


Fig. 12. Three different forgeries are taken from DEFACTO dataset [129] such as (a) Splicing, (b) Copy-Move, (c) Object removal (Inpainting), and its corresponding ground truths are in (d), (e), and (f), respectively.

region duplication in several publications (see, for example, [76, 133, 165, 183]). To avoid confusion, we refer to processing two or more photos as "cut-paste" instead of "splicing", which is the term used to refer to both copy-move and cut-paste individually. The name "erase-fill", which was coined by using the same naming strategy as copy-move and splicing, actually goes by the more well-known name "object removal", which initially referred to the application of reviving misplaced or

Table 7. The Attributes of Image Manipulation Localization Techniques Such as Copy-move, Splicing, and Object Removal

Year	Literature	Method		Fusion		Semantic Segementation	Manipulation type		
		RGB Stream	Noise Stream	When?	Method type		Copy-move	Splicing	Removal
2017	MFCN [159]	✓	-	-	-	FCN	-	✓	-
2018	RGB- N [224]	✓	SRM	Late	Bilinear pooling	Faster R-CNN	✓	✓	✓
2018	D-CNNs [166]	✓	SRM	Late	Concatenate	SLIC	✓	✓	✓
2019	H-LSTM [8]	✓	-	-	-	Patch-LSTM	✓	✓	✓
2019	ManTra- Net [197]	✓	SRM + Constrain	Early	Concatenate	Wider VGG	✓	✓	✓
2019	Den-Inecp Net [222]	✓	-	-	-	HPP	✓	-	-
2020	HP-FCN [107]	-	HPF	-	-	FCN	-	-	✓
2020	GSR-Net [223]	✓	-	-	-	Deeplabv2	✓	✓	✓
2020	CR-CNN [201]	-	Constrain	Early	Concatenate	Mask R-CNN	✓	✓	✓
2020	STRD Net [35]	✓	-	-	-	Dial-CNN and ASPP	✓	-	-
2020	SPAN [82]	✓	SRM + Constrain	Middle	Attention	Wider VGG	✓	✓	✓
2020	DisTool [12]	✓	-	late	Attention	Siamese FCN	✓	-	-
2020	MM-Net [202]	✓	Constrain	Early	Concatenate	Mask R-CNN	✓	✓	✓
2020	D-UNet [20]	✓	SRM,DWT	Middle	Concatenate	UNet	-	✓	-
2021	JPEG-ComNet [153]	✓	SRM	Middle	Concatenate	Siamese FCN	✓	✓	✓
2021	DFCN [226]	✓	-	-	-	FCN and Dilated Conv	✓	✓	✓
2021	CAT-Net [103]	✓	DCT	Late	-	HR Net	✓	✓	✓
2022	ESRNet [156]	✓	-	-	EFPN, RFBA	GRA	✓	✓	✓
2022	Edge-Enha-Transf [176]	✓	Sobel	Middle	Concatenate	Transformer	-	✓	-
2023	MVSS-Net [57]	✓	Constrain	Late	Dual attention	FCN	✓	✓	✓
2023	MSTAF [179]	✓	Patch Embedding	Late	TA Attention	Transformer	-	✓	-
2024	D-Net [205]	✓	DWT	Middle	Concatenate	Dual Encoder	-	✓	-
2024	HDF-Net [75]	✓	SRM	Late	Attention	FCN and M_ViT	✓	✓	✓
2024	IML-Net [121]	✓	Constrain	Late	Attention	Swin-Transformer	✓	✓	
2025	PIM [98]	✓	Patch Embedding	Middle	LWM	Encoder and Decoder	✓	✓	✓

damaged portions in an image [73]. To stress-fill the target region, we are now using the phrase "erase-fill". The attributes of the three categories of tampering are presented in Table 7.

The several similarities are there between these three forgeries. Firstly, all three methods of image tampering replace a portion of the original image with new graphic content, significantly altering it. Visually contrasting the modified image with the original can reveal this alteration. Second, the unaltered image is the most significant source used by copy-move and existing erase-fill applications. Third, copy-move and splicing allow more extensive source image manipulation than erase-fill. Therefore, when all the source images are displayed, it is evident that some parts have been repeated. Finally, individual category serves a specific purpose (i.e., copy-move is the most popular method for erasing an unsightly entity from a photograph by shifting it into the background). An alternate entity subtraction technique called erase-fill fills the area around the unwanted entity with texture from nearby entities. Conversely, cut-paste is typically employed to add a new entity to the unaltered image. Discarding and addition are typically add-on actions, such as adding a picturesque backdrop when an item is removed (copy-move), whereas adding an external entity is equivalent to erasing a portion of the image backdrop (erase-fill). In general, forgery such as cut-and-paste can be utilised to swap out an entity for a piece of background from another image, copy-and-move to insert new ones, and erase-and-fill to restore damaged ones.

3.7 Copy-move and Copy-paste Fingerprints

To identify such forgeries, few publications in this field focused on multiple-type or general image manipulation detection [110, 151, 159]. Most research in this field focused on identifying a single

manipulation type, such as copy-move in [46] or splicing in [214]. To accomplish this difficult task, we must reconsider the significance of the opinions about manipulation clues in real-world scenarios of image tampering. This survey offers a fresh perspective to assess current image tampering categories and their accompanying detection methodologies. By examining representative examples of image manipulation, we can see that the focus is shifting from simple to complex tampering scenarios. Edge abnormalities provide more widespread tampering signals than patch duplication by contrasting copy-move detection algorithms with splicing detection methods. Identifying the location of the manipulated region in an image becomes essential when combining tampered photographs with other modified images because detection is insufficient to demonstrate tampering. This study evaluates recently published articles and acts as a primer for newcomers to understand image manipulation and its detection. In addition, recent related works on image manipulation detection and localization are shown in Table 7. One of the most challenging tasks in multimedia forensics is the identification of **copy-move forgery** (**CMF**). This forgery aims to conceal or replicate specific details or sections of the original image. The term "copy-move forgeries" refers to copying a small portion of an image and then superimposing it on another similar image. In [119], a high-level outline for CMF identification techniques is directed. Splicing, however, involves cutting out a specific segment of an image and replacing it with another. Many works to detect CMF have been reported early, and most of them are based on either (i) Key-point-based feature matching [4, 25, 220] to find duplicate regions or (ii) Block-based feature matching [33, 89, 126, 127] to separate the image into non-overlapping sections.

However, these approaches have some issues, including their considerable computational complexity. Researchers have recently turned to data-driven approaches for identifying copy-move and cut-paste forgeries in an image [56, 119, 122, 159, 196]. Table 7 lists some methods for detecting splicing, copy-move, and inpainting manipulations, including the fusion method and segmentation network. In addition, it also lists the methodology, a summary of the article, performance characteristics, and the dataset utilized for the study's evaluation. Rao and Ni [151] describes a unique method for identifying splicing and CMF based on a deep learning architecture. Specifically, they use image manipulation procedures on the input RGB colour image to train a supervised CNN on the hierarchical characteristics. In contrast to traditional CNN, which uses a random seed to initialize its weights, this approach uses an HPF set, typically employed in estimating residuals in the spatial rich model. In this approach, the initial layer comprises weights derived from thirty HPFs, which hide the image's basic information and detect tiny artefacts created by manipulations. The autonomous feature-learning CNN architecture consists of ten separate layers. A pre-trained CNN generates dense patch-based features from the test image, and then these features are fused using the feature fusion method to get the final discriminative features. Finally, the SVM classifier uses these discriminatory features to determine whether the item's authenticity is forged. In addition, in [224], a dual-stream faster R-CNN is proposed to exploit RGB stream and noise stream features by connecting the first network to a general-purpose deep CNN. In this approach, noise stream features and RGB stream features are taken from SRM filters, and the RGB stream is shown in Figure 13.

3.8 Inpainting Fingerprints

In recent years, many deep-learning models have been proposed for image object removal (i.e., inpainting). Deep-learning-based inpainting systems can generate fully original content and attain state-of-the-art inpainting performance because they use massive datasets for visual language understanding of images. Training deep GANs for huge inpainting voids in images was first explored by Pathak et al. [140]. However, the proposed networks fail to maintain global consistency and frequently result in highly distracting visual artefacts. The generative network developed by Iizuka

Fig. 13. First column: altered image, second column: red boundary boxes in the first column, third column: local noise inconsistency between tampered and fair places, fourth column: ground truth images. Visual and noise elements work together to disclose tampering artefacts by [224].

et al. [161] supports global and local characters because of its two context discriminators. On the other hand, some works [200, 208] presented an attention mechanism that together utilises the available characteristics to evaluate the misplaced elements rather than just using the features of hidden layers. To properly manage the extensive history while preventing misuse, Wang et al. [189] proposed a multi-staged method for conducting contextual attention using images to enhance the attention mechanism. However, numerous publications [123, 209] have employed partial or gated convolutions to minimise colour disagreement and blurriness. The convolutions are hidden, renormalized, and applied only to the previously annotated area in these cases. To improve the inpainting detection performance, authors in [132] and [195] recommended employing an edge/LBP generator in the first step, followed by an image completion network in keeping with the current trend of using two-stage networks. In addition, occluded face recognition algorithms can benefit from inpainting approaches [69, 106]. In [69], inpainting-based identify-diversity GAN is introduced to enhance the capability of well-trained face recognizers in recognizing occluded faces. Furthermore, in [106], an inpainting-guided de-occlusion purification system was created for effective masked face recognition. On the flip side, various inpainting forensic techniques have been developed to combat the malicious use of inpainting modifications [2, 39, 60, 63, 83, 113, 182, 225]. They all work on the premise that forgery is more likely to occur in blocks that are highly similar to others in an image. Another effective counterfeit detection system was published by combining mapping of the core pixels, labelling the components with the largest zero-connectivity, and identifying fragment splicing in [113]. Recent work in [225] constructed a trained encoder-decoder network using a label matrix and weighted cross-entropy to record manipulation traces. These forensic methods cannot detect inpainting alterations because diffusion-based techniques do not produce visually identical blocks in the inpainted areas reported in [108]. To address this problem, the local variance of the image Laplacian along the isophote direction to detect diffusion-based inpainting is proposed in [108]. Furthermore, [197] also proposed ManTra-Net, a more general forgery localization network, which first extracts image manipulation trace features and then identifies anomalous regions by assessing how different a local feature is from its reference features. This allows for detecting complex combinations of forgeries (including inpainting). However, there are

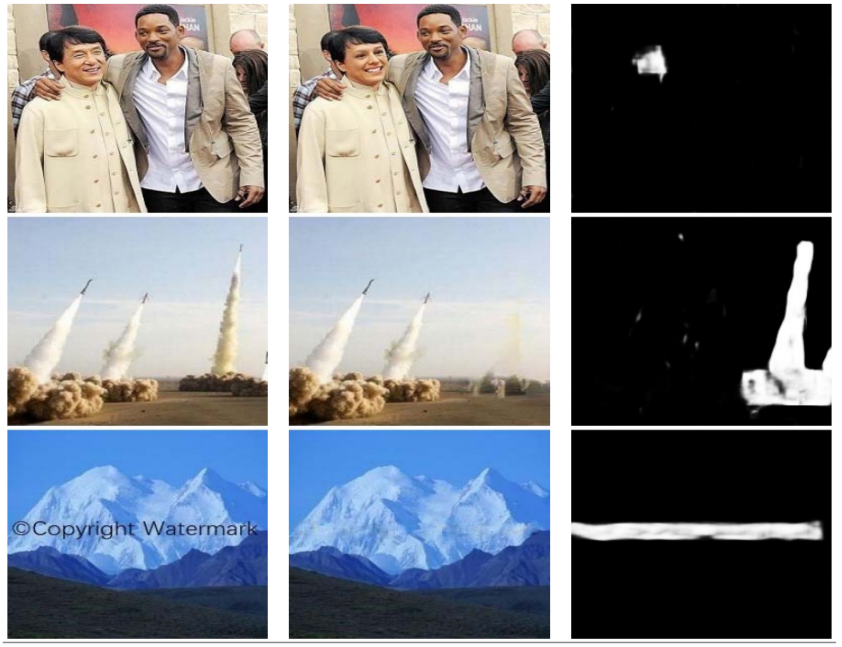


Fig. 14. First column: original image, second column: forged image created using [123, 132, 195], third column: output of IID-Net [194] with a forged image as input.

some difficult circumstances where ManTra-Net may ultimately fail. One such case is when fake features dominate the image. To identify the inpainted areas precisely, the authors in [194] present a revolutionary end-to-end **Image Inpainting Detection Network** (**IID-Net**). The enhancement, extraction, and decision blocks are utilized to construct IID-Net and indicate that many image inpainting techniques share similar artefacts. Figure 14 (c) shows IID-Net's direct detection result from Figure 14 (b) without post-processing. IID-Net was trained without using the original images in Figure 14 (a) or the inpainting methods [123, 132, 195].

Since deep learning-based inpainting techniques can utilise learned high-level semantic information to generate more complex structures and even novel objects, they may leave completely different artefacts in the inpainted regions, resulting in very poor detection performance of the aforementioned forensic methods [107]. The HP-FCN, developed by Li and Huang [107], is a DL-based approach to pinpointing the areas of an image that have been altered using deep inpainting. They employ a high-pass pre-filtering module to suppress unwanted details in the image, enhancing the contrast between inpainted and unaltered regions. Furthermore, a wavelet-based approach has demonstrated that HP-FCN can effectively detect inpainting forgeries. Despite using a training set produced with the same inpainting process and its practical significance, [107] does not explore generalizability to unseen inpainting approaches.

4 Performance Evaluation and Benchmarking

This section overviews the datasets for evaluating manipulation detection and localization techniques. On the other hand, an essential part of any quantitative investigation of a method's performance is pinpointing its limitations, and this is where the validation measures play a crucial role. For a detailed discussion of validation measures and performance measures, please refer to Appendix B (Validation Measures) and Appendix C (Performance Measures).

Table 8. Summary of Latest Image Manipulation Datasets

Year	Dataset	Image size	Images	Description
2022	MSM30K [156]	various	30,000	M1, M2, M3, M4
2020	IMD2020 [138]	various	Over 37,000	M1, M2, M3, M4
2019	GAN [130]	4256 × 256 to 1024 × 1024	Over 596,000	M1, M2, M3, M4
2019	DEFACTO[129]	480 × 640 or 640 × 480	Over 200,000	M1, M2, M3, M4, F1, F2
2019	MFC [72]	Various	Over 100,000	M1, M2, M3
2016	RTD [100]	1920 × 1080	220	M4, F1
2016	COVERAGE [193]	400 × 486	100 (authentic) + 100	M3
2015	CMH [167]	845 × 634 to 1296 × 972	108	M1, M3
2015	CVIP [6]	1000 × 700 or 700 × 1000	1,160	M3
2015	RAISE [53]	Various	8,156	F1
2015	WildWEB [213]	Various	10,646	M1, M3
2015	Wattanachote [192]	Various	2,347	F1, F2
2014	CMFDdb grip [49]	Various	3,440	F2
2013	CoMoFoD [181]	512 × 512	5,200	M3, F2
2013	MICC-F600 [3]	800 × 533 to 3888 × 2592	160	M3, F2
2013	CASIA v2.0 [58]	320 × 240 to 800 × 600	7,200 (authentic) + 5,123	M2, F1, F2
2013	CASIA v1.0 [58]	374 × 256	800 (authentic) + 921	M2, F1, F2
2012	CMEN [46]	800 × 533 to 3872 × 2592	336	M3
2012	Bianchi [22]	1024 × 1024	100	F1, F2
2011	MICC-F2000 [4]	2048 × 1536	1,300 (authentic) + 700	M2, F2
2011	MICC-F220 [4]	722 × 480 to 800 × 600	110 (authentic) + 110	M2, F2
2011	BOSSBases v0.93 [13]	512 × 512	9,074	F1
2010	Dresden [71]	Various	25,137	F1
2008	INRIA Copydays [87]	Various	1,642	F2
2006	Columbia color [80]	757 × 568 to 1, 152 × 768	183 (authentic) + 180	M2, F1
2004	Columbia gray [134]	128 × 128	933 (authentic) + 912	M2, F1
M1: Manipulation detection, M2: Splicing, M3: Copy-move, M4: Insertion, F1: TIFF format, F2: JPEG compression				

4.1 Datasets

The images that construct the tampering dataset have been altered, and the originals have been included. In the tampered image dataset, photos are typically labelled with a value of '0' for legitimate and a value of '1' for tampering. The ground truth is included in the photo alteration dataset so that it can be used to benchmark the detection. Initially, the dataset used to study image alteration was relatively small [80, 134], with only a small number of photos and a single tampering method. In the past, determining whether an image was the original or had been altered required several feature extractors to isolate the distinct traces left behind by the type of tampering, followed by comparing the predicted mask with the ground truth mask.

Having a perfect dataset that includes various tampering operations, different image formats, and photographs in varied situations is notoriously challenging. However, as deep learning has developed, which typically demands an extensive scale of the dataset [72], only a tiny handful of large datasets consisting of thousands of photos are accessible, and even these datasets are insufficient for the data-driven approaches. Consequently, investigators often create artificial images using existing datasets [8, 9, 15, 51, 90, 199]. These datasets aid in identifying image alteration in the real world by identifying the various tampering operations and pinpointing the manipulated area. Here, we surveyed a few open-source collections that investigate instances of image manipulation (see Table 8). Table 8 provides a quick overview of the image modification datasets, including details such as the year the dataset was created, the size of the photos, the number of images, and a brief explanation of the changes in the dataset.

5 Perspectives on Future Directions

To address the evolving challenges in multimedia forensics, future research must focus on advancing detection capabilities, addressing key research gaps, and leveraging multi-modal analysis. This section highlights three critical areas for future work: improving detection methods, bridging research gaps, and integrating contextual information for comprehensive analysis. As manipulation techniques become more sophisticated, the following subsections explore how forensic tools can evolve to meet these challenges effectively.

5.1 Advancing Detection Capabilities for Complex Manipulations

This survey demonstrates how much development has been made in the field of multimedia forensics during the past fifteen years. However, many problems still need fixing, new obstacles popping up daily, and a long way to go before forgery analysts reach their destination. On the other hand, there is no doubt that the introduction of data-driven methods has given a tremendous boost to both data manipulation techniques and forensic tools [191], hence enabling the exploration of novel topics. However, a more fundamental explanation of multimedia forensics is that two players exist in this field of study. The existence of competent adversaries ensures that no mechanism can provide permanent security and that novel approaches will always be required to deal with unknown threats. Given this premise, it's crucial to zero in on the most fruitful research avenues.

As the sophistication of manipulations rises, the effectiveness of any forensic tool will decline across the board. To overcome this, the ideal way to combine all the available information should be the subject of more in-depth study; therefore, numerous detection tools, networks, and methodologies must be combined. Hence, it is important to pursue multi-tool fusion and multi-asset analysis. The media assets themselves, along with any relevant proof, are becoming increasingly important to examine. Furthermore, examining a photo or video clip that was used to spread disinformation without the text, audio, and other context around it is insufficient [70, 94, 99], and [26]. While advancements in detection capabilities are crucial, addressing the technical limitations and research gaps in current forensic tools is equally important to ensure their robustness and adaptability.

5.2 Key Research Gaps in Multimedia Forensics

One of the most pressing challenges in multimedia forensics lies in the technical limitations of deep learning-based tools, particularly their inability to adapt to unseen scenarios and transformations. When we narrow our attention to tools based on deep learning, the (in)capability of deep networks to adjust to circumstances not seen during training is likely the most pressing technical issue. This problem appears in several settings. To begin with, media assets' high-order statistics, which are so valuable for forgery detection, are considerably altered by seemingly innocuous processing processes like compression, resizing, rotation, re-capturing, and so on. Malicious alterations meant to conceal forensic evidence must be taken into account alongside their innocent counterparts. All possible permutations of such transformations are unlikely to be captured by a training set. Consequently, alternate strategies should be followed to achieve greater robustness. Furthermore, deep networks need to be easily adaptable to new alterations without full re-training, which may be unattainable due to a lack of training data or entail a destructive forgetting phenomenon, to keep up with the instantaneous refinements in manipulation schemes. Addressing these research gaps paves the way for a more comprehensive approach to forensic analysis, where multi-modal and contextual information can be integrated to enhance detection accuracy.

5.3 Multi-Modal and Context-Aware Forensic Analysis

To achieve a holistic approach to multimedia forensics, future research must focus on integrating multi-modal data and contextual information alongside traditional image analysis. The scope for

future work of deep learning models for image manipulation is quite broad and promising, and some key areas include the following: Currently, deep learning models have limitations in detecting complex forms of image manipulation. There is a scope for advancement in terms of accuracy by developing new models that can better identify subtle manipulations. There is a growing interest in developing models that can incorporate multiple modalities, such as audio, video, and text, to enhance the accuracy of image manipulation identification. As deep learning models become more prevalent in image manipulation detection, attackers will likely develop adversarial attacks to bypass these models. There is a need to develop models that are robust to these attacks. Deep learning models are often viewed as "black boxes" as it can be challenging to understand how they arrived at an unavoidable decision. There is a need to develop more transparent and interpretable models to increase their accountability and trustworthiness. There is a demand for developing models that detect image manipulation in real-time, especially in security-critical applications. By addressing these challenges, the field of multimedia forensics can move closer to achieving robust, adaptable, and context-aware solutions for detecting manipulated content.

6 Conclusions

In the last two decades, only a few researchers in the research fields of law enforcement, intelligence, and private investigations have been interested in multimedia forensics. As part of manipulation detection techniques, this survey post offers numerous helpful insights, and the primary goal of this study is to provide an overview of the methods currently available in the literature for improving, segmenting, extracting, and classifying doctored images for image manipulation detection and localization. Therefore, this review has the potential to contribute to a new understanding of the evaluation of cutting-edge procedures, and it is a crucial part of the generated altered and post-processing operations. In addition, the benefits and drawbacks of the most cutting-edge methods are examined in this survey. Furthermore, definitions of performance measures such as sensitivity, specificity, accuracy, precision, F_1-score, and area AUC are included and utilized in quantitatively evaluating detection methods. More precisely, technical discussions on data-driven methods give constructive ideas for researchers to evaluate different approaches more objectively. In conclusion, while deep learning models have made great strides in image manipulation detection, much work must be done to ensure these methods are reliable, robust, and trustworthy. In addition, technical discussions, forensic applications, and potential future directions are also discussed. While this review provides a comprehensive overview of current methods and databases in digital image forensics, future work should focus on a more detailed comparative analysis of techniques to advance the field further and address its evolving challenges. This survey could be helpful to academics and forgery analysts as they settle on manipulation identification and analysis.

Acknowledgments

This research is supported by the following projects:

(1) Project titled "Deep learning applications for computer vision task" funded by NITROAA with support of Lenovo P920 and Dell Inception 7820 workstation and NVIDIA Corporation with support of NVIDIA Titan V and Quadro RTX 8000 GPU.
(2) Project titled "Computer Vision-based Smart Solutions for UAV Remote Sensing Applications through Semantic Segmentation" funded by Vishlesan I-Hub, IIT Patna, Technology Innovation Hub (NMCPS-DST), Government of India.

References

[1] Ahmed Alharbi, Hai Dong, Xun Yi, Zahir Tari, and Ibrahim Khalil. 2021. Social media identity deception detection: A survey. *ACM Computing Surveys* 54, 3 (2021), 1–35.

[2] Mohammed Aloraini, Mehdi Sharifzadeh, and Dan Schonfeld. 2020. Sequential and patch analyses for object removal video forgery detection and localization. *IEEE Transactions on Circuits and Systems for Video Technology* 31, 3 (2020), 917–930.

[3] Irene Amerini, Lamberto Ballan, Roberto Caldelli, Alberto Del Bimbo, Luca Del Tongo, and Giuseppe Serra. 2013. Copy-move forgery detection and localization by means of robust clustering with J-Linkage. *Signal Processing: Image Communication* 28, 6 (2013), 659–669.

[4] Irene Amerini, Lamberto Ballan, Roberto Caldelli, Alberto Del Bimbo, and Giuseppe Serra. 2011. A SIFT-based forensic method for copy–move attack detection and transformation recovery. *IEEE Transactions on Information Forensics and Security* 6, 3 (2011), 1099–1110.

[5] Irene Amerini, Tiberio Uricchio, Lamberto Ballan, and Roberto Caldelli. 2017. Localization of JPEG double compression through multi-domain convolutional neural networks. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, 1865–1871.

[6] Edoardo Ardizzone, Alessandro Bruno, and Giuseppe Mazzola. 2015. Copy–move forgery detection by matching triangles of keypoints. *IEEE Transactions on Information Forensics and Security* 10, 10 (2015), 2084–2094.

[7] Khosro Bahrami, Alex C. Kot, Leida Li, and Haoliang Li. 2015. Blurred image splicing localization by exposing blur type inconsistency. *IEEE Transactions on Information Forensics and Security* 10, 5 (2015), 999–1009.

[8] Jawadul H. Bappy, Cody Simons, Lakshmanan Nataraj, B. S. Manjunath, and Amit K. Roy-Chowdhury. 2019. Hybrid LSTM and encoder–decoder architecture for detection of image forgeries. *IEEE Transactions on Image Processing* 28, 7 (2019), 3286–3300.

[9] Mauro Barni, Luca Bondi, Nicolo Bonettini, Paolo Bestagini, Andrea Costanzo, Marco Maggini, Benedetta Tondi, and Stefano Tubaro. 2017. Aligned and non-aligned double JPEG detection using convolutional neural networks. *Journal of Visual Communication and Image Representation* 49, November 2017 (2017), 153–163.

[10] Mauro Barni, Andrea Costanzo, Ehsan Nowroozi, and Benedetta Tondi. 2018. CNN-based detection of generic contrast adjustment with JPEG post-processing. In *Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 3803–3807.

[11] Mauro Barni, Marco Fontani, and Benedetta Tondi. 2013. A universal attack against histogram-based image forensics. *International Journal of Digital Crime and Forensics* 5, 3 (2013), 35–52.

[12] Mauro Barni, Quoc-Tin Phan, and Benedetta Tondi. 2020. Copy-move source-target disambiguation through multi-branch CNNs. *IEEE Transactions on Information Forensics and Security* 16, 2021 (2020), 1825–1840.

[13] Patrick Bas, Tomáš Filler, and Tomáš Pevnỳ. 2011. "Break our steganographic system": The ins and outs of organizing BOSS. In *Proceedings of the 13th International Conference on Information Hiding, IH 2011, Prague, Czech Republic, May 18-20, 2011, Revised Selected Papers 13*. Springer, 59–70.

[14] Sebastiano Battiato, Oliver Giudice, and Antonino Paratore. 2016. Multimedia forensics: Discovering the history of multimedia contents. In *Proceedings of the 17th International Conference on Computer Systems and Technologies 2016*. 5–16.

[15] Belhassen Bayar and Matthew C. Stamm. 2016. A deep learning approach to universal image manipulation detection using a new convolutional layer. In *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security*. 5–10.

[16] Belhassen Bayar and Matthew C. Stamm. 2017. A generic approach towards image manipulation parameter estimation using convolutional neural networks. In *Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security*. 147–157.

[17] Belhassen Bayar and Matthew C. Stamm. 2018. Constrained convolutional neural networks: A new approach towards general purpose image manipulation detection. *IEEE Transactions on Information Forensics and Security* 13, 11 (2018), 2691–2706. DOI: https://doi.org/10.1109/TIFS.2018.2825953

[18] Sevinç Bayram, İsmail Avcıbaş, Bülent Sankur, and Nasir Memon. 2006. Image manipulation detection. *Journal of Electronic Imaging* 15, 4 (2006), 041102–041102.

[19] Alexandre Berthet and Jean-Luc Dugelay. 2020. A review of data preprocessing modules in digital image forensics methods using deep learning. In *Proceedings of the 2020 IEEE International Conference on Visual Communications and Image Processing (VCIP)*. IEEE, 281–284.

[20] Xiuli Bi, Yanbin Liu, Bin Xiao, Weisheng Li, Chi-Man Pun, Guoyin Wang, and Xinbo Gao. 2020. D-Unet: A dual-encoder u-net for image splicing forgery detection and localization. arXiv:2012.01821. Retrieved from https://arxiv.org/abs/2012.01821

[21] Tiziano Bianchi and Alessandro Piva. 2011. Detection of nonaligned double JPEG compression based on integer periodicity maps. *IEEE Transactions on Information Forensics and Security* 7, 2 (2011), 842–848.

[22] Tiziano Bianchi and Alessandro Piva. 2012. Image forgery localization via block-grained analysis of JPEG artifacts. *IEEE Transactions on Information Forensics and Security* 7, 3 (2012), 1003–1017.

[23] Tiziano Bianchi and Alessandro Piva. 2012. Reverse engineering of double JPEG compression in the presence of image resizing. In *Proceedings of the IEEE International Workshop on Information Forensics and Security (WIFS)*. 127–132.
[24] Gajanan K. Birajdar and Vijay H. Mankar. 2013. Digital image forgery detection using passive techniques: A survey. *Digital Investigation* 10, 3 (2013), 226–245.
[25] Xu Bo, Wang Junwen, Liu Guangjie, and Dai Yuewei. 2010. Image copy-move forgery detection based on SURF. In *Proceedings of the 2010 International Conference on Multimedia Information Networking and Security*. IEEE, 889–892.
[26] Christina Boididou, Symeon Papadopoulos, Markos Zampoglou, Lazaros Apostolidis, Olga Papadopoulou, and Yiannis Kompatsiaris. 2018. Detection and visualization of misleading content on twitter. *International Journal of Multimedia Information Retrieval* 7, 1 (2018), 71–86.
[27] Mehdi Boroumand and Jessica Fridrich. 2017. Scalable processing history detector for JPEG images. *Electronic Imaging* 29 (2017), 128–137.
[28] Mehdi Boroumand and Jessica J. Fridrich. 2018. Deep learning for detecting processing history of images. *Electronic Imaging* 2018 (2018), 213–1–213–9.
[29] Alanconrad Bovik. 1987. Streaking in median filtered images. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 35, 4 (1987), 493–503.
[30] Jason Bunk, Jawadul H. Bappy, Tajuddin Manhar Mohammed, Lakshmanan Nataraj, Arjuna Flenner, B. S. Manjunath, Shivkumar Chandrasekaran, Amit K Roy-Chowdhury, and Lawrence Peterson. 2017. Detection and localization of image forgeries using resampling features and deep learning. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, 1881–1889.
[31] Gang Cao, Yao Zhao, Rongrong Ni, and Xuelong Li. 2014. Contrast enhancement-based forensics in digital images. *IEEE Transactions on Information Forensics and Security* 9, 3 (2014), 515–525.
[32] Gang Cao, Antao Zhou, Xianglin Huang, Gege Song, Lifang Yang, and Yonggui Zhu. 2019. Resampling detection of recompressed images via dual-stream convolutional neural network. *Mathematical Biosciences and Engineering* 16, 5 (2019), 5022–5040. https://www.aimspress.com/article/doi/10.3934/mbe.2019253
[33] Yanjun Cao, Tiegang Gao, Li Fan, and Qunting Yang. 2012. A robust detection algorithm for copy-move forgery in digital images. *Forensic Science International* 214, 1-3 (2012), 33–43.
[34] Paola Capasso, Giuseppe Cattaneo, and Maria De Marsico. 2023. A comprehensive survey on methods for image integrity. *ACM Transactions on Multimedia Computing, Communications and Applications* 20, 11 (2023), 1–34.
[35] Beijing Chen, Weijin Tan, Gouenou Coatrieux, Yuhui Zheng, and Yun-Qing Shi. 2020. A serial image copy-move forgery localization scheme with source/target distinguishment. *IEEE Transactions on Multimedia* 23, 2021 (2020), 3506–3517.
[36] Chunhua Chen, Yun Q. Shi, and Wei Su. 2008. A machine learning based scheme for double JPEG compression detection. In *Proceedings of the2008 19th International Conference on Pattern Recognition*. IEEE, 1–4.
[37] Jiansheng Chen, Xiangui Kang, Ye Liu, and Z. Jane Wang. 2015. Median filtering forensics based on convolutional neural networ ks. *IEEE Signal Processing Letters* 22, 11 (2015), 1849–1853.
[38] Jiaxin Chen, Xin Liao, Wei Wang, and Zheng Qin. 2023. Identification of image global processing operator chain based on feature decoupling. *Information Sciences* 637, August 2023 (2023), 118961.
[39] Shengda Chen, Shunquan Tan, Bin Li, and Jiwu Huang. 2015. Automatic detection of object-based forgery in advanced video. *IEEE Transactions on Circuits and Systems for Video Technology* 26, 11 (2015), 2138–2151.
[40] Yifang Chen, Xiangui Kang, Yun Q. Shi, and Z. Jane Wang. 2019. A multi-purpose image forensic method using densely connected convolutional neural networks. *Journal of Real-Time Image Processing* 16, 01 June 2019 (2019), 725–740.
[41] Yifang Chen, Xiangui Kang, Z. Jane Wang, and Qiong Zhang. 2018. Densely connected convolutional neural network for multi-purpose image forensics under anti-forensic attacks. In *Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security*. 91–96.
[42] Yifang Chen, Zheng Wang, Z. Jane Wang, and Xiangui Kang. 2020. Automated design of neural network architectures with reinforcement learning for detection of global manipulations. *IEEE Journal of Selected Topics in Signal Processing* 14, 5 (2020), 997–1011.
[43] Yi-Lei Chen and Chiou-Ting Hsu. 2011. Detecting recompression of JPEG images via periodicity analysis of compression artifacts for tampering detection. *IEEE Transactions on Information Forensics and Security* 6, 2 (2011), 396–406.
[44] Zhipeng Chen, Yao Zhao, and Rongrong Ni. 2017. Detection of operation chain: JPEG-resampling-JPEG. *Signal Processing: Image Communication* 57, September 2017 (2017), 8–20.
[45] Zhipeng Chen, Jie Zhu, and Jun Zhang. 2022. The forensicability of operation detection in image operation chain. *IEEE Access* 10, 2022 (2022), 68557–68569.
[46] Vincent Christlein, Christian Riess, Johannes Jordan, Corinna Riess, and Elli Angelopoulou. 2012. An evaluation of popular copy-move forgery detection approaches. *IEEE Transactions on Information Forensics and Security* 7, 6 (2012), 1841–1854.

[47] Xiaoyu Chu, Yan Chen, and K. J. Ray Liu. 2015. Detectability of the order of operations: An information theoretic approach. *IEEE Transactions on Information Forensics and Security* 11, 4 (2015), 823–836.
[48] Davide Cozzolino, Diego Gragnaniello, and Luisa Verdoliva. 2014. Image forgery localization through the fusion of camera-based, feature-based and pixel-based techniques. In *Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, 5302–5306.
[49] Davide Cozzolino, Giovanni Poggi, and Luisa Verdoliva. 2014. Copy-move forgery detection based on patchmatch. In *Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, 5312–5316.
[50] Davide Cozzolino, Giovanni Poggi, and Luisa Verdoliva. 2015. Efficient dense-field copy–move forgery detection. *IEEE Transactions on Information Forensics and Security* 10, 11 (2015), 2284–2297.
[51] Davide Cozzolino, Giovanni Poggi, and Luisa Verdoliva. 2017. Recasting residual-based local descriptors as convolutional neural networks: An application to image forgery detection. In *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security*. 159–164.
[52] L. Minh Dang, Syed Ibrahim Hassan, Suhyeon Im, and Hyeonjoon Moon. 2019. Face image manipulation detection based on a convolutional neural network. *Expert Systems with Applications* 129, 1 September 2019 (2019), 156–168.
[53] Duc-Tien Dang-Nguyen, Cecilia Pasquini, Valentina Conotter, and Giulia Boato. 2015. RAISE: A raw images dataset for digital image forensics. In *Proceedings of the 6th ACM Multimedia Systems Conference*. 219–224.
[54] Alessia De Rosa, Marco Fontani, Matteo Massai, Alessandro Piva, and Mauro Barni. 2015. Second-order statistics analysis to cope with contrast enhancement counter-forensics. *IEEE Signal Processing Letters* 22, 8 (2015), 1132–1136.
[55] Cheng Deng, Zhao Li, Xinbo Gao, and Dacheng Tao. 2019. Deep multi-scale discriminative networks for double JPEG compression forensics. *ACM Transactions on Intelligent Systems and Technology* 10, 2 (2019), 1–20.
[56] Anuja Dixit and Soumen Bag. 2021. A fast technique to detect copy-move image forgery with reflection and non-affine transformation attacks. *Expert Systems with Applications* 182, 15 November 2021 (2021), 115282.
[57] Chengbo Dong, Xinru Chen, Ruohan Hu, Juan Cao, and Xirong Li. 2023. MVSS-Net: Multi-view multi-scale supervised networks for image manipulation detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 3 (2023), 3539–3553. DOI: https://doi.org/10.1109/TPAMI.2022.3180556
[58] Jing Dong, Wei Wang, and Tieniu Tan. 2013. CASIA image tampering detection evaluation database. In *Proceedings of the 2013 IEEE China Summit and International Conference on Signal and Information Processing*. IEEE, 422–426.
[59] Jing Dong, Wei Wang, Tieniu Tan, and Yun Q Shi. 2009. Run-length and edge statistics based approach for image splicing detection. In *Proceedings of the 7th International Workshop on Digital Watermarking, IWDW 2008, Busan, Korea, November 10-12, 2008*. Springer, 76–87.
[60] Luca D'Amiano, Davide Cozzolino, Giovanni Poggi, and Luisa Verdoliva. 2018. A patchmatch-based dense-field algorithm for video copy–move detection and localization. *IEEE Transactions on Circuits and Systems for Video Technology* 29, 3 (2018), 669–682.
[61] Hany Farid. 2009. Image forgery detection. *IEEE Signal Processing Magazine* 26, 2 (2009), 16–25.
[62] Hany Farid and Siwei Lyu. 2003. Higher-order wavelet statistics and their application to digital forensics. In *Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition Workshop*. IEEE, 94–94.
[63] Chunhui Feng, Zhengquan Xu, Shan Jia, Wenting Zhang, and Yanyan Xu. 2016. Motion-adaptive frame deletion detection for digital video forensics. *IEEE Transactions on Circuits and Systems for video Technology* 27, 12 (2016), 2543–2554.
[64] Xiaoying Feng, Ingemar J. Cox, and Gwenaël Doërr. 2011. An energy-based method for the forensic detection of re-sampled images. In *Proceedings of the IEEE International Conference on Multimedia and Expo*. 1–6.
[65] William D. Ferreira, Cristiane B. R. Ferreira, Gelson da Cruz Júnior, and Fabrizzio Soares. 2020. A review of digital image forensics. *Computers and Electrical Engineering* 85, July 2020 (2020), 106685.
[66] A. Jessica Fridrich, B. David Soukal, and A. Jan Lukáš. 2003. Detection of copy-move forgery in digital images. In *Proceedings of the Digital Forensic Research Workshop*. Citeseer.
[67] Jessica Fridrich and Jan Kodovsky. 2012. Rich models for steganalysis of digital images. *IEEE Transactions on Information Forensics and Security* 7, 3 (2012), 868–882.
[68] Andrew C. Gallagher. 2005. Detection of linear and cubic interpolation in JPEG compressed images. In *Proceedings of the Canadian Conference on Computer and Robot Vision (CRV'05)*. IEEE, 65–72.
[69] Shiming Ge, Chenyu Li, Shengwei Zhao, and Dan Zeng. 2020. Occluded face recognition in the wild by identity-diversity inpainting. *IEEE Transactions on Circuits and Systems for Video Technology* 30, 10 (2020), 3387–3397.
[70] Shiming Ge, Fanzhao Lin, Chenyu Li, Daichi Zhang, Weiping Wang, and Dan Zeng. 2022. Deepfake video detection via predictive representation learning. *ACM Transactions on Multimedia Computing, Communications, and Applications* 18, 2s (2022), 1–21.
[71] Thomas Gloe and Rainer Böhme. 2010. The 'dresden image database' for benchmarking digital image forensics. In *Proceedings of the 2010 ACM Symposium on Applied Computing*. 1584–1590.

[72] Haiying Guan, Mark Kozak, Eric Robertson, Yooyoung Lee, Amy N. Yates, Andrew Delgado, Daniel Zhou, Timothee Kheyrkhah, Jeff Smith, and Jonathan Fiscus. 2019. MFC datasets: Large-scale benchmark datasets for media forensic challenge evaluation. In *Proceedings of the 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW)*. IEEE, 63–72.

[73] Christine Guillemot and Olivier Le Meur. 2013. Image inpainting: Overview and recent advances. *IEEE Signal Processing Magazine* 31, 1 (2013), 127–144.

[74] Abhinav Gupta and Divya Singhal. 2019. Global median filtering forensic method based on pearson parameter statistics. *IET Image Processing* 13, 12 (2019), 2045–2057.

[75] Ruidong Han, Xiaofeng Wang, Ningning Bai, Yihang Wang, Jianpeng Hou, and Jianru Xue. 2024. HDF-Net: Capturing homogeny difference features to localize the tampered image. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46, 12 (2024), 10005–10020.

[76] Junfeng He, Zhouchen Lin, Lifeng Wang, and Xiaoou Tang. 2006. Detecting doctored JPEG images via DCT coefficient analysis. In *Proceedings of the 9th European Conference on Computer Vision–ECCV 2006, Graz, Austria, May 7-13, 2006*. Springer, 423–435.

[77] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 770–778. DOI: https://doi.org/10.1109/CVPR.2016.90

[78] Zhongwei He, Wei Lu, Wei Sun, and Jiwu Huang. 2012. Digital image splicing detection based on markov features in DCT and DWT domain. *Pattern Recognition* 45, 12 (2012), 4292–4299.

[79] Jeff Heaton. 2018. Ian goodfellow, yoshua bengio, and aaron courville: Deep learning: The MIT press, 2016, 800 pp, ISBN: 0262035618. *Genetic Programming and Evolvable Machines* 19, 1-2 (2018), 305–307.

[80] Yu-Feng Hsu and Shih-Fu Chang. 2006. Detecting image splicing using geometry invariants and camera characteristics consistency. In *Proceedings of the 2006 IEEE International Conference on Multimedia and Expo*. IEEE, 549–552.

[81] Yu-Feng Hsu and Shih-Fu Chang. 2010. Camera response functions for image forensics: An automatic algorithm for splicing detection. *IEEE Transactions on Information Forensics and Security* 5, 4 (2010), 816–825.

[82] Xuefeng Hu, Zhihan Zhang, Zhenye Jiang, Syomantak Chaudhuri, Zhenheng Yang, and Ram Nevatia. 2020. SPAN: Spatial pyramid attention network for image manipulation localization. In *Proceedings of the European Conference on Computer Vision*. Springer, 312–328.

[83] Fangjun Huang, Xiaochao Qu, Hyoung Joong Kim, and Jiwu Huang. 2015. Reversible data hiding in JPEG images. *IEEE Transactions on Circuits and Systems for Video Technology* 26, 9 (2015), 1610–1621.

[84] Xiaosa Huang, Shilin Wang, and Gongshen Liu. 2018. Detecting double JPEG compression with same quantization matrix based on dense CNN feature. In *Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 3813–3817.

[85] Israr Hussain, Shunquan Tan, Bin Li, Xinghong Qin, Dostdar Hussain, and Jiwu Huang. 2021. A novel deep learning framework for double JPEG compression detection of small size blocks. *Journal of Visual Communication and Image Representation* 80, October 2021 (2021), 103269.

[86] Tukey J. 1971. Exploratory data analysis. MA: Addison-wesley. *Wesley* 23, 4 (1971), 413–414.

[87] Herve Jegou, Matthijs Douze, and Cordelia Schmid. 2008. Hamming embedding and weak geometry consistency for large scale image search-extended version. In *Proceedings of the 10th European Conference on Computer Vision–ECCV 2008, Marseille, France, October 12-18, 2008*. Springer.

[88] Xinghao Jiang, Peisong He, Tanfeng Sun, Feng Xie, and Shilin Wang. 2017. Detection of double compression with the same coding parameters based on quality degradation mechanism analysis. *IEEE Transactions on Information Forensics and Security* 13, 1 (2017), 170–185.

[89] Zhao Junhong. 2010. Detection of copy-move forgery based on one improved LLE method. In *Proceedings of the 2010 2nd International Conference on Advanced Computer Control*. IEEE, 547–550.

[90] Vijayakumar Kadha and Santos Kumar Das. 2022. Robust first quality factor estimation for double compressed and resized images. In *Proceedings of the 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*. 1550–1554. DOI: https://doi.org/10.1109/ICOEI53556.2022.9776910

[91] Vijayakumar Kadha, V. V. N. J. Sri Lakshmi Nandikattu, Sambit Bakshi, and Santos Kumar Das. 2023. Forensic analysis of manipulation chains: A deep residual network for detecting JPEG-manipulation-JPEG. *Forensic Science International: Digital Investigation* 47, December 2023 (2023), 301623.

[92] Vijaya Kumar Kadha, Prashant Deshmukh, Krishna Chaitanya Rayasam, and Santos Kumar Das. 2022. Robust manipulation detection scheme for post-JPEG compressed images using CNN. In *Proceedings of the 2022 IEEE 19th India Council International Conference (INDICON)*. 1–6. DOI: https://doi.org/10.1109/INDICON56171.2022.10040157

[93] Xiangui Kang, Matthew C. Stamm, Anjie Peng, and K. J. Ray Liu. 2013. Robust median filtering forensics using an autoregressive model. *IEEE Transactions on Information Forensics and Security* 8, 9 (2013), 1456–1468.

[94] Dhruv Khattar, Jaipal Singh Goud, Manish Gupta, and Vasudeva Varma. 2019. MVAE: Multimodal variational autoencoder for fake news detection. In *Proceedings of the World Wide Web Conference*. 2915–2921.
[95] Matthias Kirchner. 2008. Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue. In *Proceedings of the ACM Workshop on Multimedia and Security*. 11–20.
[96] Matthias Kirchner and Rainer Bohme. 2008. Hiding traces of resampling in digital images. *IEEE Transactions on Information Forensics and Security* 3, 4 (2008), 582–592.
[97] Matthias Kirchner and Thomas Gloe. 2009. On resampling detection in re-compressed images. In *Proceedings of the IEEE International Workshop on Information Forensics and Security (WIFS)*. 21–25.
[98] Chenqi Kong, Anwei Luo, Shiqi Wang, Haoliang Li, Anderson Rocha, and Alex C. Kot. 2025. Pixel-inconsistency modeling for image manipulation localization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 99 (2025), 1–18.
[99] Pavel Korshunov, Michael Halstead, Diego Castan, Martin Graciarena, Mitchell McLaren, Brian Burns, Aaron Lawson, and Sebastien Marcel. 2019. Tampered speaker inconsistency detection with phonetically aware audio-visual features. In *Proceedings of the International Conference on Machine Learning*.
[100] Paweł Korus and Jiwu Huang. 2016. Multi-scale analysis strategies in PRNU-based tampering localization. *IEEE Transactions on Information Forensics and Security* 12, 4 (2016), 809–824.
[101] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. In *Proceedings of the International Conference on Neural Information Processing Systems—Volume 1 (NIPS'12)*. Curran Associates Inc., Red Hook, NY, USA, 1097–1105.
[102] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2017. ImageNet classification with deep convolutional neural networks. *Communications of the ACM* 60, 6 (2017), 84–90.
[103] Myung-Joon Kwon, In-Jae Yu, Seung-Hun Nam, and Heung-Kyu Lee. 2021. CAT-Net: Compression artifact tracing network for detection and localization of image splicing. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 375–384.
[104] Bin Li, Hu Luo, Haoxin Zhang, Shunquan Tan, and Zhongzhou Ji. 2017. A multi-branch convolutional neural network for detecting double JPEG compression. arXiv:1710.05477. Retrieved from https://arxiv.org/abs/1710.05477
[105] Bin Li, Haoxin Zhang, Hu Luo, and Shunquan Tan. 2019. Detecting double JPEG compression and its related anti-forensic operations with CNN. *Multimedia Tools and Applications* 78, April 2019 (2019), 8577–8601.
[106] Chenyu Li, Shiming Ge, Daichi Zhang, and Jia Li. 2020. Look through masks: Towards masked face recognition with de-occlusion distillation. In *Proceedings of the 28th ACM International Conference on Multimedia*. 3016–3024.
[107] Haodong Li and Jiwu Huang. 2019. Localization of deep inpainting using high-pass fully convolutional network. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 8301–8310.
[108] Haodong Li, Weiqi Luo, and Jiwu Huang. 2017. Localization of diffusion-based inpainting in digital images. *IEEE Transactions on Information Forensics and Security* 12, 12 (2017), 3050–3064.
[109] Haodong Li, Weiqi Luo, Xiaoqing Qiu, and Jiwu Huang. 2016. Identification of various image operations using residual-based features. *IEEE Transactions on Circuits and Systems for Video Technology* 28, 1 (2016), 31–45.
[110] Haodong Li, Weiqi Luo, Xiaoqing Qiu, and Jiwu Huang. 2017. Image forgery localization via integrating tampering possibility maps. *IEEE Transactions on Information Forensics and Security* 12, 5 (2017), 1240–1252.
[111] Yuanman Li, Jiaxiang You, Jiantao Zhou, Wei Wang, Xin Liao, and Xia Li. 2022. Image operation chain detection with machine translation framework. *IEEE Transactions on Multimedia* 25, 2023 (2022), 6852–6867.
[112] Yaohua Liang, Yanmei Fang, Shangjun Luo, and Bing Chen. 2019. Image resampling detection based on convolutional neural network. In *Proceedings of the 2019 15th International Conference on Computational Intelligence and Security (CIS)*. IEEE, 257–261.
[113] Zaoshan Liang, Gaobo Yang, Xiangling Ding, and Leida Li. 2015. An efficient forgery detection algorithm for object removal by exemplar-based image inpainting. *Journal of Visual Communication and Image Representation* 30, July 2015 (2015), 75–85.
[114] Guang-Yu Liao, T. Nodes, and N. Gallagher. 1985. Output distributions of two-dimensional median filters. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 33, 5 (1985), 1280–1295.
[115] Xin Liao, Kaide Li, Xinshan Zhu, and K. J. Ray Liu. 2020. Robust detection of image operator chain with two-stream convolutional neural network. *IEEE Journal of Selected Topics in Signal Processing* 14, 5 (2020), 955–968.
[116] Xufeng Lin, Chang-Tsun Li, and Yongjian Hu. 2013. Exposing image forgery through the detection of contrast enhancement. In *Proceedings of the 2013 IEEE International Conference on Image Processing*. IEEE, 4467–4471.
[117] Zhouchen Lin, Rongrong Wang, Xiaoou Tang, and Heung-Yeung Shum. 2005. Detecting doctored images using camera response normality and consistency. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. IEEE, 1087–1092.
[118] Anan Liu, Zhengyu Zhao, Chengqian Zhang, and Yuting Su. 2017. Median filtering forensics in digital images based on frequency-domain features. *Multimedia Tools and Applications* 76, November 2017 (2017), 22119–22132.

[119] Bo Liu and Chi-Man Pun. 2018. Locating splicing forgery by fully convolutional networks and conditional random field. *Signal Processing: Image Communication* 66, August 2018 (2018), 103–112.
[120] Chang Liu and Matthias Kirchner. 2019. CNN-based rescaling factor estimation. In *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security*. 119–124.
[121] Xuntao Liu, Yuzhou Yang, Haoyue Wang, Qichao Ying, Zhenxing Qian, Xinpeng Zhang, and Sheng Li. 2024. Multi-view feature extraction via tunable prompts is enough for image manipulation localization. In *Proceedings of the 32nd ACM International Conference on Multimedia*. 9999–10007.
[122] Yaqi Liu, Qingxiao Guan, and Xianfeng Zhao. 2018. Copy-move forgery detection based on convolutional kernel network. *Multimedia Tools and Applications* 77, July 2018 (2018), 18269–18293.
[123] Yaqi Liu, Qingxiao Guan, Xianfeng Zhao, and Yun Cao. 2018. Image forgery localization based on multi-scale convolutional neural networks. In *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security*. 85–90.
[124] Shenghai Luo, Anjie Peng, Hui Zeng, Xiangui Kang, and Li Liu. 2019. Deep residual learning using data augmentation for median filtering forensics of digital images. *IEEE Access* 7, 2019 (2019), 80614–80621.
[125] Weiqi Luo, Zhenhua Qu, Jiwu Huang, and Guoping Qiu. 2007. A novel method for detecting cropped and re-compressed image block. In *Proceedings of the 2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*. IEEE, II–217.
[126] Gavin Lynch, Frank Y. Shih, and Hong-Yuan Mark Liao. 2013. An efficient expanding block algorithm for image copy-move forgery detection. *Information Sciences* 239, 1 August 2013 (2013), 253–265.
[127] Babak Mahdian and Stanislav Saic. 2007. Detection of copy–move forgery using a method based on blur moment invariants. *Forensic Science International* 171, 2-3 (2007), 180–189.
[128] Babak Mahdian and Stanislav Saic. 2008. Blind authentication using periodic properties of interpolation. *IEEE Transactions on Information Forensics and Security* 3, 3 (2008), 529–538.
[129] Gaël Mahfoudi, Badr Tajini, Florent Retraint, Frédéric Morain-Nicolier, Jean Luc Dugelay, and Marc Pic. 2019. DEFACTO: Image and face manipulation dataset. In *Proceedings of the 2019 27th European Signal Processing Conference (EUSIPCO)*. 1–5. DOI: https://doi.org/10.23919/EUSIPCO.2019.8903181
[130] Francesco Marra, Diego Gragnaniello, Luisa Verdoliva, and Giovanni Poggi. 2019. Do GANs leave artificial fingerprints?. In *Proceedings of the 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. IEEE, 506–511.
[131] Aniruddha Mazumdar and Prabin Kumar Bora. 2020. Siamese convolutional neural network-based approach towards universal image forensics. *IET Image Processing* 14, 13 (2020), 3105–3116. DOI: https://doi.org/10.1049/iet-ipr.2019.1114
[132] Kamyar Nazeri, Eric Ng, Tony Joseph, Faisal Qureshi, and Mehran Ebrahimi. 2019. Edgeconnect: Structure Guided image inpainting using edge prediction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*.
[133] Tian-Tsong Ng, Shih-Fu Chang, and Qibin Sun. 2004. Blind detection of photomontage using higher order statistics. In *Proceedings of the 2004 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, V–V.
[134] Tian-Tsong Ng, Shih-Fu Chang, and Q Sun. 2004. A data set of authentic and spliced image blocks. *Columbia University, ADVENT Technical Report* 4 (2004), 2033–2004.
[135] Yakun Niu, Xiaolong Li, Yao Zhao, and Rongrong Ni. 2021. Detection of double JPEG compression with the same quantization matrix via convergence analysis. *IEEE Transactions on Circuits and Systems for Video Technology* 32, 5 (2021), 3279–3290.
[136] Yakun Niu, Lei Tan, Lei Zhang, and Xianyu Zuo. 2024. TMFNet: Two-stream multi-channels fusion networks for color image operation chain detection. arXiv:2409.07701. Retrieved from https://arxiv.org/abs/2409.07701
[137] Yakun Niu, Yao Zhao, and Rongrong Ni. 2017. Robust median filtering detection based on local difference descriptor. *Signal Processing: Image Communication* 53, April 2017 (2017), 65–72.
[138] Adam Novozamsky, Babak Mahdian, and Stanislav Saic. 2020. IMD2020: A large-scale annotated dataset tailored for detecting manipulated images. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision Workshops*. 71–80.
[139] Jinseok Park, Donghyeon Cho, Wonhyuk Ahn, and Heung-Kyu Lee. 2018. Double JPEG detection in mixed JPEG quality factors using deep convolutional neural network. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 636–652.
[140] Deepak Pathak, Philipp Krahenbuhl, Jeff Donahue, Trevor Darrell, and Alexei A. Efros. 2016. Context encoders: Feature learning by inpainting. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2536–2544.
[141] Tomas Pevny and Jessica Fridrich. 2008. Detection of double-compression in JPEG images for applications in steganography. *IEEE Transactions on Information Forensics and Security* 3, 2 (2008), 247–258.

[142] Stefan Pfennig and Matthias Kirchner. 2012. Spectral methods to determine the exact scaling factor of resampled digital images. In *Proceedings of the International Symposium on Communications, Control and Signal Processing*. IEEE, 1–6.

[143] Alessandro Piva. 2013. An overview on image forensics. *ISRN Signal Processing* 2013, 1 (2013), 496701.

[144] Alin C. Popescu and Hany Farid. 2004. Statistical tools for digital forensics. In *Proceedings of the International Workshop on Information Hiding*. Springer, 128–147.

[145] Alin C. Popescu and Hany Farid. 2005. Exposing digital forgeries by detecting traces of resampling. *IEEE Transactions on Signal Processing* 53, 2 (2005), 758–767.

[146] Tong Qiao, Ran Shi, Xiangyang Luo, Ming Xu, Ning Zheng, and Yiming Wu. 2019. Statistical model-based detector via texture weight map: Application in re-sampling authentication. *IEEE Transactions on Multimedia* 21, 5 (2019), 1077–1092. DOI: https://doi.org/10.1109/TMM.2018.2872863

[147] Tong Qiao, Aichun Zhu, and Florent Retraint. 2018. Exposing image resampling forgery by using linear parametric model. *Multimedia Tools and Applications* 77, January 2018 (2018), 1501–1523.

[148] Zhenhua Qu, Weiqi Luo, and Jiwu Huang. 2008. A convolutive mixing model for shifted double JPEG compression with application to passive image authentication. In *Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 1661–1664.

[149] Muhammad Ali Qureshi and Mohamed Deriche. 2015. A bibliography of pixel-based blind image forgery detection techniques. *Signal Processing: Image Communication* 39, Part A November 2015 (2015), 46–74.

[150] Kapil Rana, Gurinder Singh, and Puneet Goyal. 2022. MSRD-CNN: Multi-scale residual deep CNN for general-purpose image manipulation detection. *IEEE Access* 10, 2022 (2022), 41267–41275. DOI: https://doi.org/10.1109/ACCESS.2022.3167714

[151] Yuan Rao and Jiangqun Ni. 2016. A deep learning approach to detection of splicing and copy-move forgeries in images. In *Proceedings of the 2016 IEEE International Workshop on Information Forensics and Security (WIFS)*. IEEE, 1–6.

[152] Yuan Rao and Jiangqun Ni. 2016. A deep learning approach to detection of splicing and copy-move forgeries in images. In *Proceedings of the IEEE International Workshop on Information Forensics and Security (WIFS)*. 1–6. DOI: https://doi.org/10.1109/WIFS.2016.7823911

[153] Yuan Rao and Jiangqun Ni. 2021. Self-supervised domain adaptation for forgery localization of JPEG compressed images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 15034–15043.

[154] Hareesh Ravi, A. Venkata Subramanyam, and Sabu Emmanuel. 2015. ACE–an effective anti-forensic contrast enhancement technique. *IEEE Signal Processing Letters* 23, 2 (2015), 212–216.

[155] Judith A. Redi, Wiem Taktak, and Jean-Luc Dugelay. 2011. Digital image forensics: A booklet for beginners. *Multimedia Tools and Applications* 51, January 2011 (2011), 133–162.

[156] Ruyong Ren, Shaozhang Niu, Hua Ren, Shubin Zhang, Tengyue Han, and Xiaohai Tong. 2022. ESRNet: Efficient search and recognition network for image manipulation detection. *ACM Transactions on Multimedia Computing, Communications, and Applications* 18, 4 (2022), 1–23.

[157] Anderson Rocha, Walter Scheirer, Terrance Boult, and Siome Goldenstein. 2011. Vision of the unseen: Current trends and challenges in digital image and video forensics. *ACM Computing Surveys* 43, 4 (2011), 1–42.

[158] Seung-Jin Ryu, Matthias Kirchner, Min-Jeong Lee, and Heung-Kyu Lee. 2013. Rotation invariant localization of duplicated image regions based on zernike moments. *IEEE Transactions on Information Forensics and Security* 8, 8 (2013), 1355–1370.

[159] Ronald Salloum, Yuzhuo Ren, and C-C Jay Kuo. 2018. Image splicing localization using a multi-task fully convolutional network (MFCN). *Journal of Visual Communication and Image Representation* 51, February 2018 (2018), 201–209.

[160] Priyanka Samanta and Shweta Jain. 2024. SmartHash: Perceptual hashing for image tampering detection and authentication. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*. 1983–1993.

[161] Iizuka Satoshi, Simo-Serra Edgar, and Ishikawa Hiroshi. 2017. Globally and locally consistent image completion. *ACM Transactions on Graphics* 36, 4 (2017), 3073659.

[162] Victor Schetinger, Massimo Iuliani, Alessandro Piva, and Manuel M. Oliveira. 2017. Image forgery detection confronts image composition. *Computers and Graphics* 68, November 2017 (2017), 152–163.

[163] Victor Schetinger, Manuel M. Oliveira, Roberto da Silva, and Tiago J. Carvalho. 2017. Humans are easily fooled by digital images. *Computers and Graphics* 68, November 2017 (2017), 142–151.

[164] Wuyang Shan, Yaohua Yi, Junying Qiu, and Aiguo Yin. 2019. Robust median filtering forensics using image deblocking and filtered residual fusion. *IEEE Access* 7, 2019 (2019), 17174–17183.

[165] Yun Q. Shi, Chunhua Chen, and Wen Chen. 2007. A natural image model approach to splicing detection. In *Proceedings of the 9th Workshop on Multimedia and Security*. 51–62.

[166] Zenan Shi, Xuanjing Shen, Hui Kang, and Yingda Lv. 2018. Image manipulation detection and localization based on the dual-domain convolutional neural networks. *IEEE Access* 6, 2018 (2018), 76437–76453.
[167] Ewerton Silva, Tiago Carvalho, Anselmo Ferreira, and Anderson Rocha. 2015. Going deeper into copy-move forgery detection: Exploring image telltales via multi-scale analysis and voting processes. *Journal of Visual Communication and Image Representation* 29, May 2015 (2015), 16–32.
[168] Karen Simonyan and Andrew Zisserman. 2015. Very deep convolutional networks for large-scale image recognition. In *Proceedings of the International Conference on Learning Representations*. Yoshua Bengio and Yann LeCun (Eds.), http://arxiv.org/abs/1409.1556
[169] Gurinder Singh, Kapil Rana, Puneet Goyal, and Sathish Kumar. 2024. NVMS-Net: A novel constrained noise-view multi-scale network for detecting general image processing based manipulations. *IEEE Transactions on Artificial Intelligence* 6, 5 (2024), 1233–1247.
[170] Divya Singhal, Abhinav Gupta, Anurag Tripathi, and Ravi Kothari. 2020. CNN-based multiple manipulation detector using frequency domain features of image residuals. *ACM Transactions on Intelligent Systems and Technology* 11, 4 (2020), 1–26.
[171] K. Sitara and Babu M. Mehtre. 2016. Digital video tampering detection: An overview of passive techniques. *Digital Investigation* 18 (2016), 8–22.
[172] Matthew Stamm and K. J. Ray Liu. 2008. Blind forensics of contrast enhancement in digital images. In *Proceedings of the 2008 15th IEEE International Conference on Image Processing*. IEEE, 3112–3115.
[173] Matthew C. Stamm and K. J. Ray Liu. 2010. Forensic estimation and reconstruction of a contrast enhancement mapping. In *Proceedings of the 2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 1698–1701.
[174] Matthew C. Stamm, Min Wu, and K. J. Ray Liu. 2013. Information forensics: An overview of the first decade. *IEEE Access* 1, 2023 (2013), 167–200.
[175] Jee-Young Sun, Seung-Wook Kim, Sang-Won Lee, and Sung-Jea Ko. 2018. A novel contrast enhancement forensics based on convolutional neural networks. *Signal Processing: Image Communication* 63, April 2018 (2018), 149–160.
[176] Yu Sun, Rongrong Ni, and Yao Zhao. 2022. ET: Edge-enhanced transformer for image splicing detection. *IEEE Signal Processing Letters* 29, 2022 (2022), 1232–1236.
[177] Ali Taimori, Farbod Razzazi, Alireza Behrad, Ali Ahmadi, and Massoud Babaie-Zadeh. 2016. Quantization-unaware double JPEG compression detection. *Journal of Mathematical Imaging and Vision* 54, March 2016 (2016), 269–286.
[178] Ali Taimori, Farbod Razzazi, Alireza Behrad, Ali Ahmadi, and Massoud Babaie-Zadeh. 2021. A part-level learning strategy for JPEG image recompression detection. *Multimedia Tools and Applications* 80, March 2021 (2021), 12235–12247.
[179] Yuxuan Tan, Yuanman Li, Limin Zeng, Jiaxiong Ye, Wei Wang, and Xia Li. 2023. Multi-scale target-aware framework for constrained splicing detection and localization. In *Proceedings of the 31st ACM International Conference on Multimedia*. 8790–8798.
[180] Hongshen Tang, Rongrong Ni, Yao Zhao, and Xiaolong Li. 2018. Median filtering detection of small-size image based on CNN. *Journal of Visual Communication and Image Representation* 51, February 2018 (2018), 162–168.
[181] Dijana Tralic, Ivan Zupancic, Sonja Grgic, and Mislav Grgic. 2013. CoMoFoD—New database for copy-move forgery detection. In *Proceedings of the International Symposium on Electronics in Marine (ELMAR)-2013*. IEEE, 49–54.
[182] DANG Thanh Trung, Azeddine Beghdadi, and Mohamed-Chaker Larabi. 2014. Blind inpainting forgery detection. In *Proceedings of the 2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 1019–1023.
[183] Diaa M. Uliyan, Hamid A. Jalab, Ainuddin W. Abdul Wahab, Palaiahnakote Shivakumara, and Somayeh Sadeghi. 2016. A novel forged blurred region detection system for image forensic applications. *Expert Systems with Applications* 64, 1 December 2016 (2016), 1–10.
[184] David Vázquez-Padín, Pedro Comesana, and Fernando Pérez-González. 2015. An SVD approach to forensic image resampling detection. In *Proceedings of the European Signal Processing Conference (EUSIPCO)*. IEEE, 2067–2071.
[185] David Vazquez-Padin, Fernando Pérez-González, and Pedro Comesana-Alfaro. 2017. A random matrix approach to the forensic analysis of upscaled images. *IEEE Transactions on Information Forensics and Security* 12, 9 (2017), 2115–2130.
[186] Dong-ping Wang, Tiegang Gao, and Fusheng Yang. 2018. A forensic algorithm against median filtering based on coefficients of image blocks in frequency domain. *Multimedia Tools and Applications* 77, September 2018 (2018), 23411–23427.
[187] Jinwei Wang, Qiye Ni, Yang Zhang, Xiangyang Luo, Yunqing Shi, Jiangtao Zhai, and Sunil Kr Jha. 2020. Median filtering detection based on quaternion convolutional neural network. *Comput Mater Continua* 65, 1 (2020), 929–943.
[188] Menglu Wang, Xueyang Fu, Jiawei Liu, and Zheng-Jun Zha. 2022. JPEG compression-aware image forgery localization. In *Proceedings of the 30th ACM International Conference on Multimedia*. 5871–5879.

[189] Ning Wang, Jingyuan Li, Lefei Zhang, and Bo Du. 2019. MUSICAL: Multi-scale image contextual attention learning for inpainting. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. 3748–3754.
[190] Qing Wang and Rong Zhang. 2016. Double JPEG compression forensics based on a convolutional neural network. *EURASIP Journal on Information Security* 2016, 1 (2016), 1–12.
[191] Tianyi Wang, Harry Cheng, Kam Pui Chow, and Liqiang Nie. 2023. Deep convolutional pooling transformer for deepfake detection. *ACM Transactions on Multimedia Computing, Communications and Applications* 19, 6 (2023), 1–20.
[192] Kanoksak Wattanachote, Timothy K. Shih, Wen-Lung Chang, and Hon-Hang Chang. 2015. Tamper detection of JPEG image due to seam modifications. *IEEE Transactions on Information Forensics and Security* 10, 12 (2015), 2477–2491.
[193] Bihan Wen, Ye Zhu, Ramanathan Subramanian, Tian-Tsong Ng, Xuanjing Shen, and Stefan Winkler. 2016. COVERAGE—A novel database for copy-move forgery detection. In *Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 161–165.
[194] Haiwei Wu and Jiantao Zhou. 2021. IID-Net: Image inpainting detection network via neural architecture search and attention. *IEEE Transactions on Circuits and Systems for Video Technology* 32, 3 (2021), 1172–1185.
[195] Haiwei Wu, Jiantao Zhou, and Yuanman Li. 2021. Deep generative model for image inpainting with local binary pattern learning and spatial attention. *IEEE Transactions on Multimedia* 24, 2022 (2021), 4016–4027.
[196] Yue Wu, Wael Abd-Almageed, and Prem Natarajan. 2018. Image copy-move forgery detection via an end-to-end deep neural network. In *Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 1907–1915.
[197] Yue Wu, Wael AbdAlmageed, and Premkumar Natarajan. 2019. Mantra-net: Manipulation tracing network for detection and localization of image forgeries with anomalous features. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 9543–9552.
[198] Ankit Yadav and Dinesh Kumar Vishwakarma. 2023. MRT-Net: Auto-adaptive weighting of manipulation residuals and texture clues for face manipulation detection. *Expert Systems with Applications* 232, 1 December 2023 (2023), 120898.
[199] Yanyang Yan, Wenqi Ren, and Xiaochun Cao. 2018. Recolored image detection via a deep discriminative model. *IEEE Transactions on Information Forensics and Security* 14, 1 (2018), 5–17.
[200] Zhaoyi Yan, Xiaoming Li, Mu Li, Wangmeng Zuo, and Shiguang Shan. 2018. Shift-net: Image inpainting via deep feature rearrangement. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 1–17.
[201] Chao Yang, Huizhou Li, Fangting Lin, Bin Jiang, and Hao Zhao. 2020. Constrained R-CNN: A general image manipulation detection model. In *Proceedings of the 2020 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 1–6.
[202] Chao Yang, Zhiyu Wang, Huawei Shen, Huizhou Li, and Bin Jiang. 2021. Multi-modality image manipulation detection. In *Proceedings of the 2021 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 1–6.
[203] Lisha Yang, Pengpeng Yang, Rongrong Ni, and Yao Zhao. 2020. Xception-based general forensic method on small-size images. In *Proceedings of the Advances in Intelligent Information Hiding and Multimedia Signal Processing*. Springer, 361–369.
[204] Lisha Yang, Pengpeng Yang, Rongrong Ni, and Yao Zhao. 2020. Xception-based general forensic method on small-size images. In *Advances in Intelligent Information Hiding and Multimedia Signal Processing: Proceedings of the 15th International Conference on IIH-MSP in conjunction with the 12th International Conference on FITAT, July 18–20, Jilin, China, Volume 2*. Springer, 361–369.
[205] Zonglin Yang, Bo Liu, Xiuli Bi, Bin Xiao, Weisheng Li, Guoyin Wang, and Xinbo Gao. 2024. D-Net: A dual-encoder network for image splicing forgery detection and localization. *Pattern Recognition* 155, November 2024 (2024), 110727.
[206] Jiaxiang You, Yuanman Li, Jiantao Zhou, Zhongyun Hua, Weiwei Sun, and Xia Li. 2021. A transformer based approach for image manipulation chain detection. In *Proceedings of the 29th ACM International Conference on Multimedia*. 3510–3517.
[207] In-Jae Yu, Seung-Hun Nam, Wonhyuk Ahn, Myung-Joon Kwon, and Heung-Kyu Lee. 2020. Manipulation classification for JPEG images using multi-domain features. *IEEE Access* 8, 2020 (2020), 210837–210854. DOI: https://doi.org/10.1109/ACCESS.2020.3037735
[208] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S. Huang. 2018. Generative image inpainting with contextual attention. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5505–5514.
[209] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S. Huang. 2019. Free-form image inpainting with gated convolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 4471–4480.
[210] Luo Yu, Yujin Zhang, Hua Han, Lijun Zhang, and Fei Wu. 2019. Robust median filtering forensics by CNN-based multiple residuals learning. *IEEE Access* 7, 2019 (2019), 120594–120602.
[211] Yang Yu, Rongrong Ni, Wenjie Li, and Yao Zhao. 2022. Detection of AI-manipulated fake faces via mining generalized features. *ACM Transactions on Multimedia Computing, Communications, and Applications* 18, 4 (2022), 1–23.

[212] Hai-Dong Yuan. 2011. Blind forensics of median filtering in digital images. *IEEE Transactions on Information Forensics and Security* 6, 4 (2011), 1335–1345.
[213] Markos Zampoglou, Symeon Papadopoulos, and Yiannis Kompatsiaris. 2015. Detecting image splicing in the wild (web). In *Proceedings of the 2015 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. IEEE, 1–6.
[214] Markos Zampoglou, Symeon Papadopoulos, and Yiannis Kompatsiaris. 2017. Large-scale evaluation of splicing localization algorithms for web images. *Multimedia Tools and Applications* 76, 4 (2017), 4801–4834.
[215] Cong Zhang, Dawei Du, Lipeng Ke, Honggang Qi, and Siwei Lyu. 2018. Global contrast enhancement detection via deep multi-path network. In *Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE, 2815–2820.
[216] Jun Zhang, Yixin Liao, Xinshan Zhu, Hongquan Wang, and Jie Ding. 2020. A deep learning approach in the discrete cosine transform domain to median filtering forensics. *IEEE Signal Processing Letters* 27, 2020 (2020), 276–280.
[217] Ying Zhang, Jonathan Goh, Lei Lei Win, and Vrizlynn L. L. Thing. 2016. Image region forgery detection: A deep learning approach. *SG-CRC* 2016, 14 (2016), 1–11.
[218] Yujin Zhang, Shenghong Li, Shilin Wang, and Yun Qing Shi. 2014. Revealing the traces of median filtering using high-order local ternary patterns. *IEEE Signal Processing Letters* 21, 3 (2014), 275–279.
[219] Xudong Zhao, Shilin Wang, Shenghong Li, Jianhua Li, and Quanqiao Yuan. 2013. Image splicing detection based on noncausal markov model. In *Proceedings of the 2013 IEEE International Conference on Image Processing*. IEEE, 4462–4466.
[220] J. Zheng and L. Chang. 2014. Detection technology of tampering image based on harris corner points. *Journal of Computer Information Systems* 10, 2014 (2014), 1481–1488.
[221] Lilei Zheng, Ying Zhang, and Vrizlynn LL Thing. 2019. A survey on image tampering and its detection in real-world photos. *Journal of Visual Communication and Image Representation* 58, January 2019 (2019), 380–399.
[222] Jun-Liu Zhong and Chi-Man Pun. 2019. An end-to-end dense-inception net for image copy-move forgery detection. *IEEE Transactions on Information Forensics and Security* 15, 2020 (2019), 2134–2146.
[223] Peng Zhou, Bor-Chun Chen, Xintong Han, Mahyar Najibi, Abhinav Shrivastava, Ser-Nam Lim, and Larry Davis. 2020. Generate, segment, and refine: Towards generic manipulation segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 13058–13065.
[224] Peng Zhou, Xintong Han, Vlad I. Morariu, and Larry S. Davis. 2018. Learning rich features for image manipulation detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1053–1061.
[225] Xinshan Zhu, Yongjun Qian, Xianfeng Zhao, Biao Sun, and Ya Sun. 2018. A deep learning approach to patch-based image inpainting forensics. *Signal Processing: Image Communication* 67, September 2018 (2018), 90–99.
[226] Peiyu Zhuang, Haodong Li, Shunquan Tan, Bin Li, and Jiwu Huang. 2021. Image tampering localization using a dense fully convolutional network. *IEEE Transactions on Information Forensics and Security* 16 (2021), 2986–2999.
[227] Hao Zou, Pengpeng Yang, Rongrong Ni, and Yao Zhao. 2021. Anti-forensics of image contrast enhancement based on generative adversarial network. *Security and Communication Networks* 2021 (2021), 1–8.

Received 20 March 2024; revised 23 March 2025; accepted 1 April 2025