Multimedia Tools and Applications (2021) 80:6247–6310
https://doi.org/10.1007/s11042-020-09974-4

A comprehensive survey on passive techniques for digital video forgery detection

Nitin Arvind Shelke[1] · Singara Singh Kasana[1]

Received: 8 February 2020 / Revised: 6 August 2020 / Accepted: 24 September 2020 /
Published online: 14 October 2020

Abstract

Digital videos are one of the most widespread forms of multimedia in day to day life. These are widely transferred over social networking websites such as Facebook, Instagram, WhatsApp, YouTube, etc. through the Internet. Availability of modern and easy to use editing tools have facilitated the modification of the contents of the digital videos. Therefore, it has become an essential concern for the legitimacy, trustworthiness, and authenticity of these digital videos. Digital video forgery detection aims to identify the manipulations in the video and to check its authenticity. These techniques can be divided into active and passive techniques. In this paper, a comprehensive survey on video forgery detection using passive techniques have been presented. The primary goal of this survey is to study and analyze the existing passive video forgery detection techniques. Firstly, the preliminary information required for understanding video forgery detection is presented. Later, a brief survey of existing passive video forgery detection techniques based on the features, forgery identified, datasets used, and performance parameters detail along with their limitations are reviewed. Then, anti-forensics strategy and deepfake detection in the video are discussed. After that, standard benchmark video forgery datasets and the generalized architecture for passive video forgery detection techniques are discussed. Finally, few open challenges in the field of passive video forgery detection are also described.

Keywords Video forgery detection · Inter-frame forgery · Intra-frame forgery · Passive techniques · Video anti-forensics · Deepfake detection

1 Introduction

Digital video is an ordered collection of images captured by a digital camera. It also contains audio and other data. People are becoming heavily dependent on multimedia contents in day

✉ Nitin Arvind Shelke
nshelke_phd17@thapar.edu

Singara Singh Kasana
singara@thapar.edu

[1] Thapar Institute of Engineering and Technology, Patiala, Punjab, India

to day life, particularly on digital videos. The surveillance camera is also one of the treasures of contemporary technology used at offices, homes, and various public places have gained enormous popularity as an efficient safety measure. It's been the fact that video footages are treated as proof in most of the nations against the sort of crimes. Also, due to easy access to advanced editing software and use of the latest smartphone, it is easy for anyone to perform the manipulations in digital video and falsify it. The intentional modification made in the digital video for falsification is called a video forgery, and it may be hard for human beings to decide the authenticity of those digital videos by the naked eye. Hence, it becomes essential to analyze and decide if video content is original or modified in order to use as a piece of evidence in court. Digital forgery detection techniques are therefore needed to inspect the integrity and authenticity of digital videos.

The digital video forgery detection is a process to validate whether the digital video contents have undergone any intentional manipulation. The techniques to detect the forgery in a digital video can be generally categorized as active and passive. Active techniques use pre-embedded information such as watermark or signature to check the integrity and authenticity of a video. In contrast, passive techniques work in the absence of pre-embedded data. But in most of the cases, videos do not contain pre-embedded information such as watermark or signature, in that case, it is tough to detect the manipulation using an active approach. So, In recent years, passive video forgery detection techniques are getting considerable attention in the scientific community, as depicted in Fig. 1. It shows the pictorial representation of the publications on video forgery detection using passive techniques over the last 15 years (i.e., from 2006 to 2020). The selection process of the papers is based on Query Firing. The keywords such as *video forgery detection* and *video forgery* are used to fire the query on standard digital libraries such as IEEE, Springer, and Elsevier.

Some of the surveys on video forgery detection have already been published: Rocha et al. [89], Wahab et al. [127], Pandey et al. [84], Sitara et al. [106], Mizher et al. [75], Singh et al. [104], Johnston et al. [47]. It has been observed in the mentioned surveys that

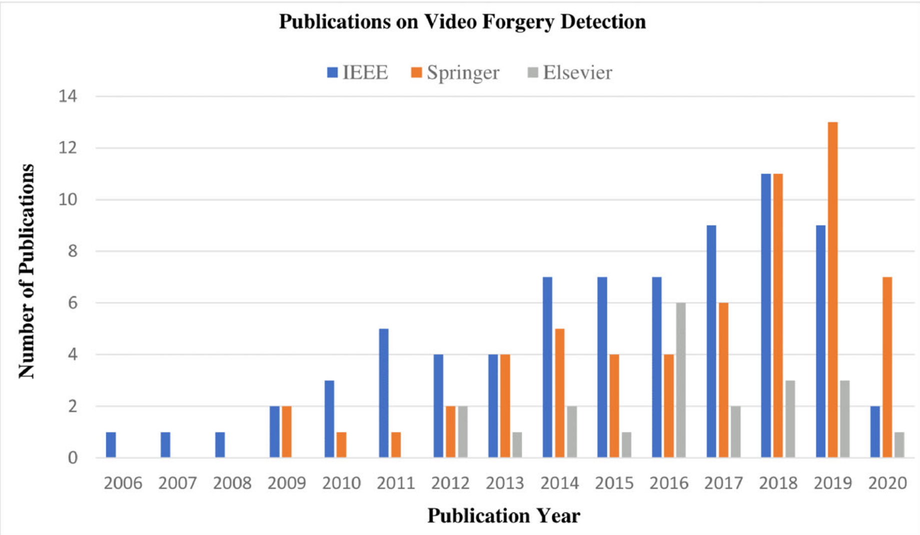

Fig. 1 Publications over the last 15 years on video forgery detection using passive techniques

1) Critical explanation about the topic is missing. 2) Systematic, easy to understand, and comprehensive survey on passive video forgery detection techniques are not done yet. 3) Deepfake detection in the video is not discussed in any of the survey paper. 4) All Performance parameters used for testing and validation of technique are not described thoroughly. 5) Discussion on standard benchmark datasets for video forgery has not done. 6) Provide limited research paths for future directions.

Our survey is different from the surveys as mentioned above in a way that a systematic method is followed in order to perform an exhaustive study on video forgery detection and delivers the in-depth literature for passive video forgery detection techniques categorized based on the feature or method used. The major highlights of this study can be precisely given, as follows,

- Basic terminology related to video forgery detection is introduced.
- The systematic and detailed survey on a passive video forgery detection technique is presented.
- Anti-forensics strategies and deepfake detection in the video are also discussed.
- The standard benchmark video forgery datasets are overviewed.
- The generalized architecture of passive video forgery detection is presented.
- Hopeful challenges and future research direction in the passive video forgery detection are also discussed.

The paper is catalogued as follows. Section 2 deals with the basic terminology required for understanding video forgery detection. Section 3 gives a detailed survey of existing passive video forgery detection techniques. Sections 4 and 5 address the anti-forensic strategies and deepfake detection in videos, respectively. Section 6 focuses on the detailed analysis of existing benchmark video forgery datasets. Section 7 presents the generalized architecture design for passive video forgery detection. Section 8 illustrates the discussion and new challenges in passive video forgery detection. Section 9 covers the conclusions.

2 Basic terminology in digital video forgery

This section presents the basic terminologies need to understand this survey.

2.1 Types of video forgeries

There are several types of forgery present in the digital video, pretty commonly divided into two subcategories, such as intra-frame forgery and inter-frame forgery. These forgeries can be performed using video editing tools such as Adobe Premiere Pro, Adobe Photoshop, *etc*. Figure 2 shows the types of digital video forgeries.

2.1.1 Intra-frame forgery

In this type of forgery, the original contents of particular frames are manipulated. It is also called as spatial based video tampering. Some of the intra-frame forgery types are as follows.

a) *Copy-Move Forgery:* It is one of the most common types of forgery performed on digital image/video [62]. In this type of forgery, an attacker can insert or delete an object from a video scene. At the same time, it can be used for creating duplicate objects in the

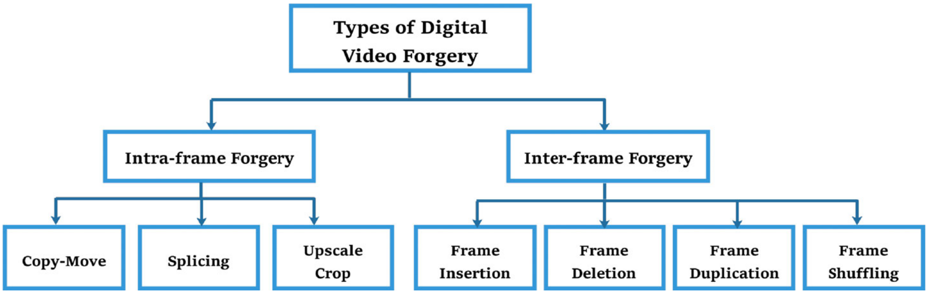

Fig. 2 Types of digital video forgeries

video by copying a portion of the video frame and pasting it to another location either in the same or the different frame of the video. Therefore, it is also called as copy-paste forgery or region manipulation forgery. The operations performed in copy-move forgery can also be used for hiding the desired area in the frame [18, 30]. Figure 3. shows an example of copy-move forgery in the video, wherein (a) part, the frame region (a flower) is copied and pasted to the other place in the same video frame (i.e., new object is created into the video frame). And in (b) part, a keyboard is removed from the actual video frame, which is highlighted by a yellow curve. Copy-move forgery is also called it as inpainting forgery which is used for removing certain objects from digital images or videos and fill that area with matching background content. Inpainting can be done in one of two ways:

– Temporal Copy and Paste Impainting: In Temporal Copy and Paste (TCP) inpainting, forged area filled-up using similar pixels from the adjacent regions of the same

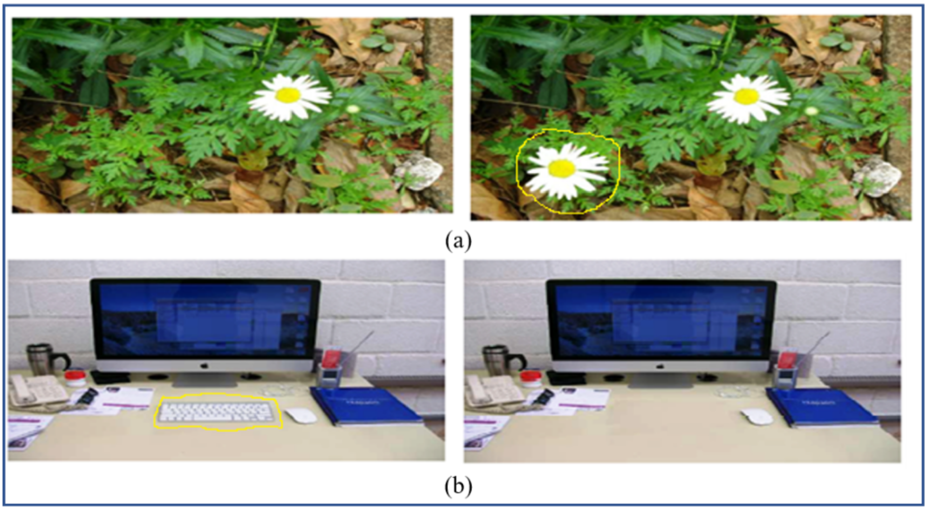

Fig. 3 Copy-move forgery in Video **a** Frame region (a flower) is copied and pasted it to another place **b** A keyboard is removed from the actual video frame which is marked by a yellow curve (also, called as video inpainting forgery)

video frame or with the help of the most coherent blocks from the frames adjacent to the affected frames.
- Exemplar Based Texture Synthesis Impainting: In Exemplar Based Texture Synthesis (ETS) inpainting, the missing areas of a video frame are filled with the use of sample textures.

b) *Splicing:* In this type of forgery, a new video frame is formed by photocopying a piece from one video frame and pasting it to another one. Figure 4 shows an example of splicing forgery in a video in which new composed video frame is formed by merging the object of two video frames.

c) *Upscale Crop:* The outer part of the video frame is crop out in upscale crop to remove some region or object [102]. Figure 5 shows an example of upscale crop forgery wherein (a) shows the original video frame and (b) shows the frame after performing upscale forgery (a walking lady is removed).

2.1.2 Inter-frame forgery

These types of forgery alter the order of frames in a video in some of the other ways. Figure 6 shows the inter-frame forgeries in digital video. It is also called as temporal tampering. The various types of inter-frame forgery are as follows.

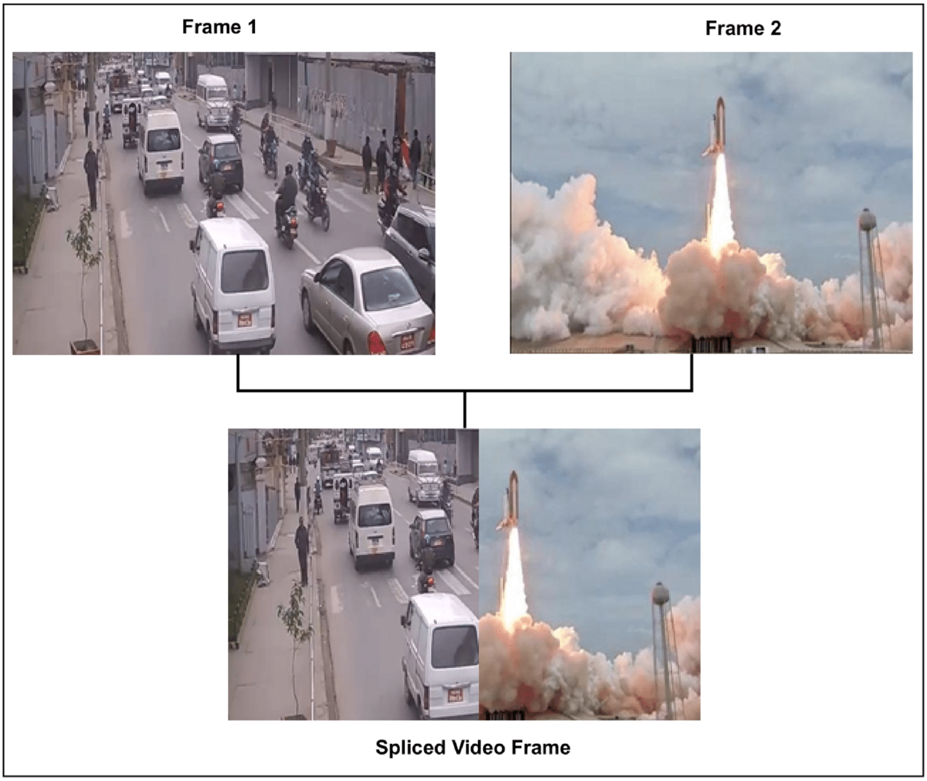

Fig. 4 Splicing forgery in video (two different frames are merged into a single frame)

Fig. 5 Upscale crop forgery in video **a**) Original video frame **b**) Frame after upscale crop forgery (a walking lady is removed)

a) *Frame Deletion:* This type of manipulation purposefully removes some of the frames in a video to produce false evidence as an unlawful activity. Figure 7 shows the frame deletion forgery in the video wherein part (a) is an original video sequence, and part (b) shows the forged video sequence after performing the frame deletion forgery in which the third and fourth frame is deleted from the original video sequence.

b) *Frame Duplication:*

This type of forgery intentionally duplicates some of the frames in a video. Figure 8 shows the frame duplication forgery in the video wherein part (a) is an original video sequence and part (b) shows the forged video sequence after performing the frame duplication forgery in which the sixth frame is duplicated in the place of the third frame.

Frame-mirroring is one of the form frame duplication forgery mentioned in [122], which copies a some of the frames from the input video and pastes its mirrored copy in the same video at some random locations. Frame mirroring is shown in Fig. 9 wherein (a) part shows the original video sequence and (b) part shows the forged video sequence created after performing frame mirroring forgery where the mirrored copy of 2nd frame is copied and pasted at location 2 denoted by M2 whereas as a mirrored copy of 6th frame is copied and pasted at location 5 denoted by M6.

c) *Frame Insertion:* In the frame insertion forgery, frames from other videos or the same video are added intentionally at some random position for any illegal activity or fake evidence. Figure 10 shows the frame insertion forgery in the video wherein part (a) is an original video sequence and part (b) shows the forged video sequence after performing insertion forgery in which frame I1 and I2 from another video are added in between the 2nd and 3rd frame of the original video sequence.

d) *Frame Shuffling/Replication:* This forgery shuffles or alters the original order of video frames, which gives the different meaning to the original video. Figure 11 shows the frame shuffling forgery in the video in which some of the frames in an original video sequence are shuffled wherein (a) part is an original video sequence and (b) part shows the forged video sequence after performing the frame shuffling forgery wherein 4th frame is shuffled with 2nd frame.

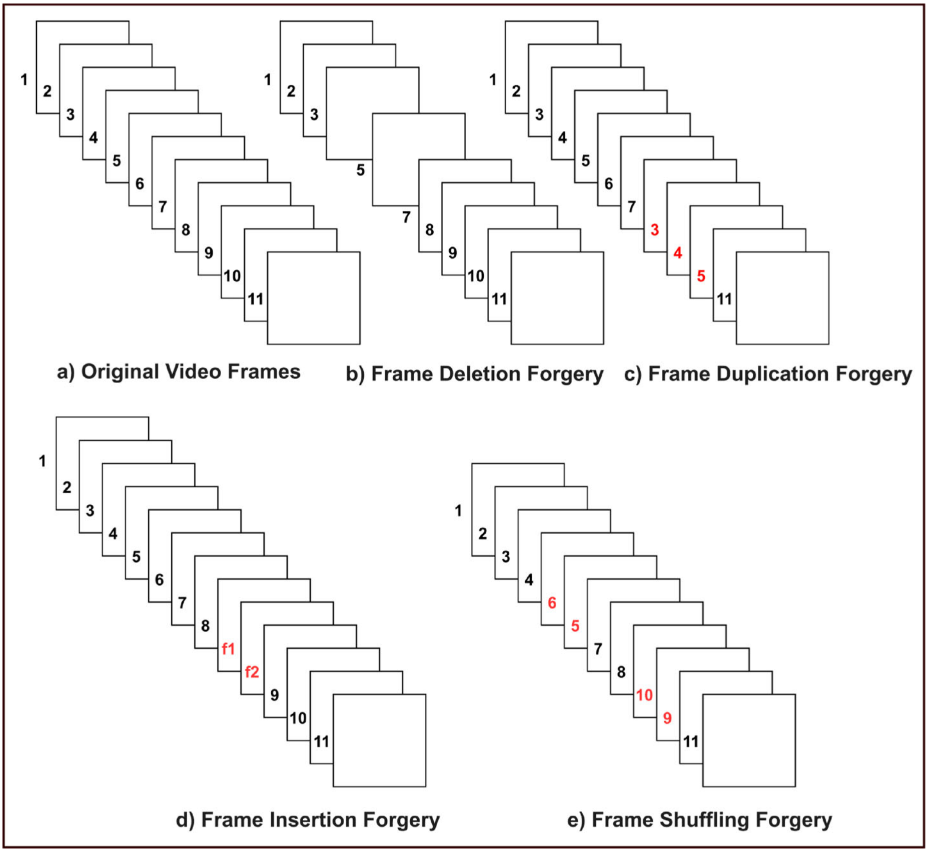

Fig. 6 Inter-frame forgeries in the video **a** Represent the original video sequence **b** Frame 4 and 6 is deleted from the original video sequence **c** Frames 3, 4 & 5 (marked by *red* color) are duplicated **d** Frame f1 & f2 is inserted into an original video sequence **e** Frames 5, 6 and 9, 10 (marked by *red* color) are shuffled

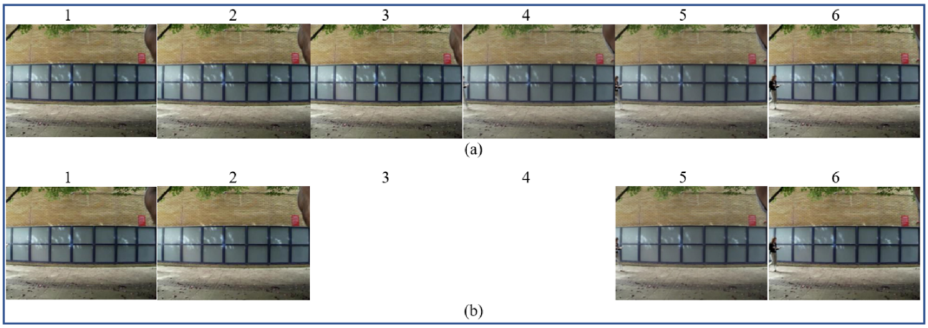

Fig. 7 Frame deletion forgery **a** Original video sequence **b** Forged video sequence after deletion forgery (3rd and 4th frame is deleted from the video sequence)

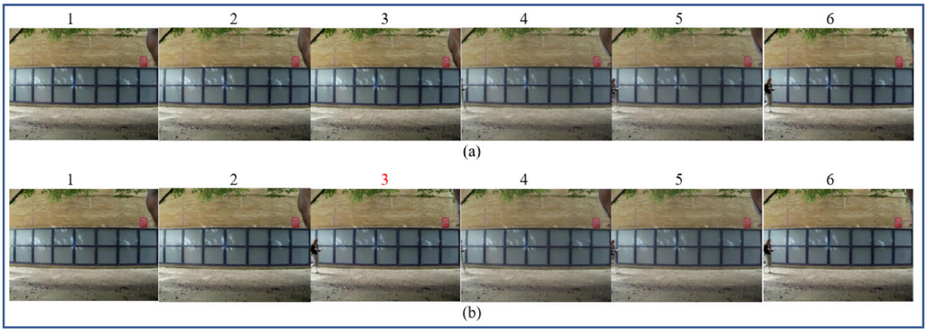

Fig. 8 Frame duplication forgery **a** Original video sequence **b** Forged video sequence after performing the duplication forgery (6th frame is duplicated in place of 3rd frame)

2.2 Performance parameters

To evaluates the performance of digital video forgery detection techniques, the common measures used by the different authors are mentioned in this section.

$$PR = \frac{TP}{TP + FP} \tag{1}$$

$$RR = \frac{TP}{TP + FN} \tag{2}$$

$$TNR = \frac{TN}{TN + FP} \tag{3}$$

$$FPR = \frac{FP}{FP + TN} \tag{4}$$

$$MR = \frac{FP + FN}{TP + FN + TN + FP} \tag{5}$$

$$DA = \frac{TP + TN}{TP + FN + TN + FP} \tag{6}$$

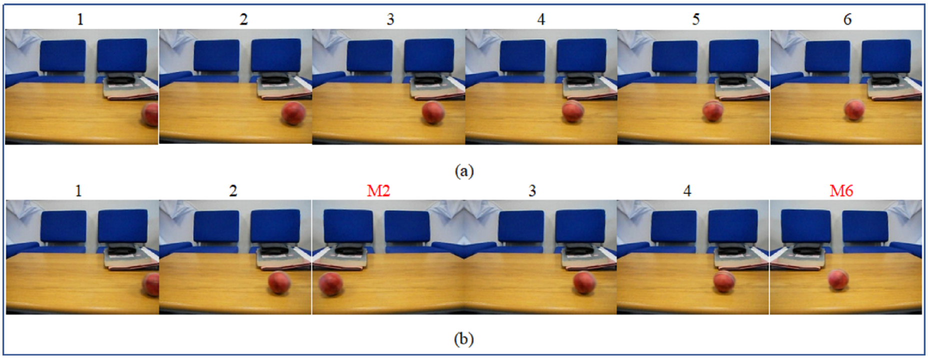

Fig. 9 Frame Mirroring forgery **a** Original video sequence **b** Forged video sequence after performing mirroring forgery

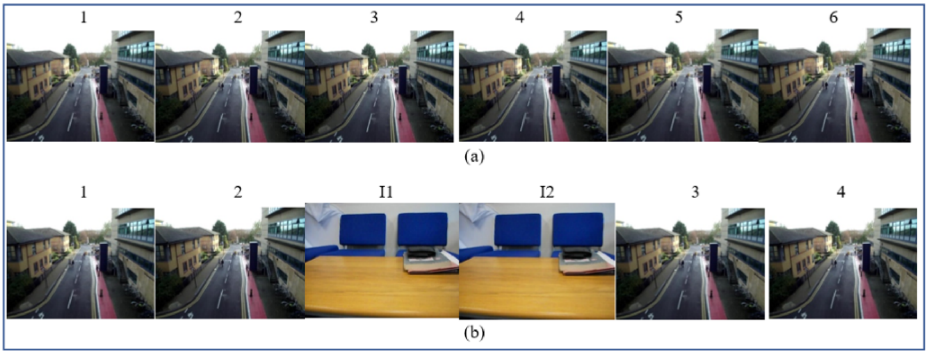

Fig. 10 Frame insertion forgery **a** Original video sequence **b** Forged video sequence after insertion forgery (I1 and I2 frames is added in between 2nd and 3rd frame)

$$F1Score = 2 \times \frac{RR \times PR}{RR + PR} \tag{7}$$

$$PFACC = \frac{Correctly\_classified\_pristine\_frames}{Pristine\_frames} \tag{8}$$

$$FFACC = \frac{Correctly\_classified\_forged\_frames}{Forged\_frames} \tag{9}$$

$$DFACC = \frac{Correctly\_classified\_double\_compressed\_frames}{double\_compressed\_frames} \tag{10}$$

$$FACC = \frac{Correctly\_classified\_frames}{All\_the\_frames} \tag{11}$$

$$VACC = \frac{Correctly\_classified\_video\_clips}{All\_the\_video\_clips} \tag{12}$$

True Positive is given by TP, which is the count of genuine video frames that are categorized as authentic i.e., correct positive detection. False Negative is given by FN, which is the count of forged video frames that are categorized as authentic i.e., incorrect negative detection. True Negative is given by TN, which is the count of forged video frames

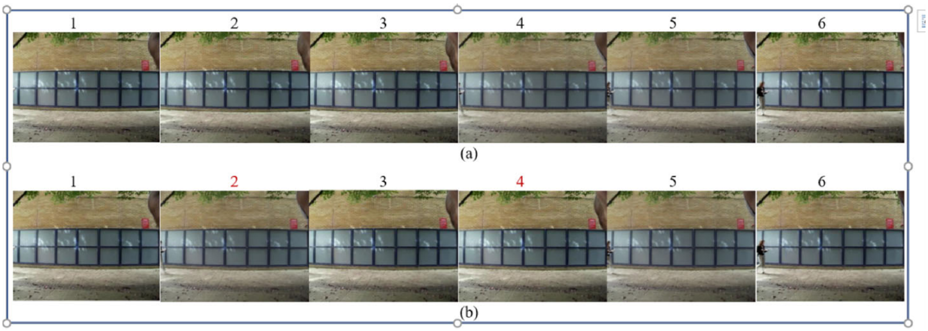

Fig. 11 Frame Shuffling Forgery **a** Original video sequence **b** Forged video sequence after performing replication forgery (4th frame is shuffled with 2nd frame)

that are categorized as forged i.e., correct negative detection. False Positive is given by FP, which is the count of genuine video frames that are categorized as forged i.e., incorrect positive detection. PR denotes Precision Rate, is computed as the number of correct positive detections divided by the total number of positive detections. RR denotes Recall Rate also, called as Sensitivity (SN) or True Positive Rate (TPR), is computed as the number of correct positive detections divided by the total number of positives. TNR denotes the True Negative Rate, also called as Specificity (SP), is computed as the number of correct negative detections divided by the total number of negatives. FPR denotes False Positive Rate, is computed as the number of incorrect positive detections divided by the total number of negatives. FPR can also be calculated as 1−TNR. MR denotes Misclassification Rate, also called as Error Rate, is calculated as the number of all incorrect detections divided by the total number of sample present in the dataset. DA denotes Detection Accuracy is computed as the number of all correct detections divided by the total number of samples in the dataset. F1 Score is a weighted average or harmonic mean of Recall and Precision. Apart from the above-mentioned parameters, the Pristine Frame Accuracy (PFACC), Forged Frame accuracy (FFACC), Double-compressed Frame Accuracy (DFACC), Frame Accuracy (FACC) and Video Accuracy (VACC) are the parameters defined by Chen et al. [20]. PFACC is the ratio of correctly classified original frames to all the original frames. FFACC is the ratio of correctly classified forged frames to all the forged frames. DFACC is the ratio of correctly classified double compressed frame to all the double compressed frames. FACC is the ratio of correctly classified frames to all the available frames (forged as well as original). VACC is the correctly classified videos to all the videos. Receiver Operating Characteristics (ROC) curve is one of the parameters which is used to plot the fraction of TP *vs* FP.

3 Video forgery detection techniques

The techniques for the detection of the forgery in a digital video can be generally categorized as active and passive. The main aim of this section is to study passive techniques designed for video forgery detection.

3.1 Active techniques

In these techniques, authentication information such as watermark or signature is inserted in a digital video that enables the authenticity and integrity of its contents [97]. If someone has manipulated the content of a video, then the watermark or signature embedded in the video is getting changed that gives the clear indication that video has been manipulated [96]. The advantage of active techniques is that forgery detection in the video is straightforward due to the presence of information like a watermark or signature. But in most of the cases, videos downloaded over the Internet do not contain a watermark or signature, in that case, it is tough to detect the manipulation. The limitation of these techniques is that if the videos do not contain pre-embedded information like watermark or signature, then it is not possible to detect the manipulation. Another issue is that it reduces the quality of an original video due to the presence of embedded information.

3.2 Passive techniques

Passive techniques depend on the internal characteristics of the digital video itself instead of information that provide to check the originality of video. Passive techniques work in the absence of pre-embedded data such as watermark or signature to check the integrity and authenticity of a video. Without knowing about the pre-embedded information inside the video, it becomes a challenging task for the researcher to work on passive techniques. Hence in recent years, passive techniques on video forgery detection have become noteworthy attention in the scientific community. Passive digital video forgery detection techniques investigate the artifacts left after the forgeries to distinguish the original videos with the tampered ones. The passive techniques are alternatively called as blind techniques as it works under the assumption that forgeries produce certain kind of static and temporal artifacts in a video which is to be checked for identifying the manipulated videos. Figure 12 shows the categorization of passive video forgery detection techniques on the basis features/artifacts used.

3.2.1 Compression artifacts based techniques

Digital videos are generally compressed through MPEG-1, MPEG-2, MPEG-3, MPEG-4 and H-264 coding standard to optimize the storage space and transmission time. Compression artifacts-based techniques used the coding clues or artifacts acquired during the process of compression to detect the forgery present in the video. The compression artifacts used in video forgery detection is shown in Fig. 13.

The manipulations in digital videos are performed in the uncompressed domain. To perform the forgery in a video, someone must decode it first, make changes and then recompress it which we generally called as double compression. The Compression artifacts look at the specific characteristics of video such as compression properties, variations in the quantization parameters after double compression, periodic features, variations in the Discrete Cosine Transform (DCT) coefficients, and properties of GOPs (Group of Pictures). Thus, the existing compression shall expose the forgery in the video. In compression artifact techniques, GOP's analysis plays a crucial part in the detection of falsification in the video. The GOP term is related to MPEG compressed video. Figure 14 shows the structure of the GOP in the video. The frames in GOP's are arranged in a specific order such as intra-frame (I), predictive frame (P) and bi-directionally predictive frame (B), each having a varying degree of compression [102]. I frames are called as intra-coded frames or independent frames and need a lot of data storage and offer the least compression ratio. Whereas, P frames are known as predicted, or dependent frames that contain only information that is distinct from it's previous I or P frame, and it requires less space as compared to I frame. During encoding, frames in a video are grouped in GOP's according to a structure that begins with an

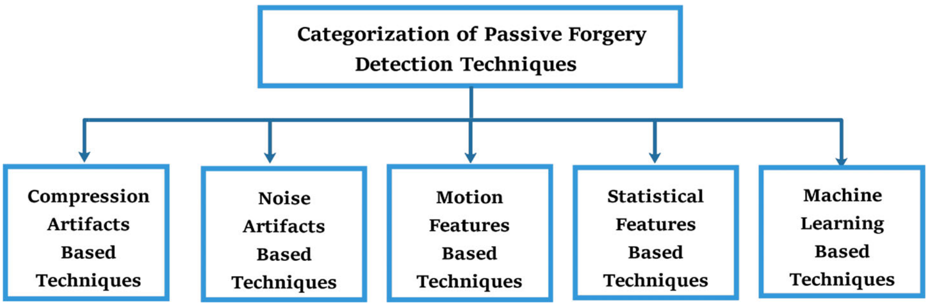

Fig. 12 Categorization of passive video forgery detection techniques

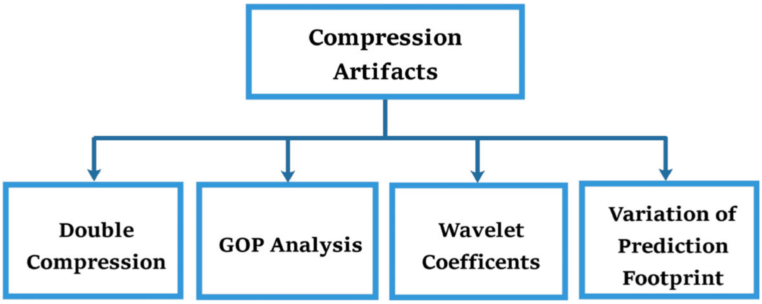

Fig. 13 Compression artifacts used in video forgery detection

I-frame and then allows a number of P and B frames [52]. Table 1 shows the analysis of video forgery detection techniques based on compression artifacts.

Wang et al. [130] focused on MPEG compressed videos and explained the fact that static and temporal features are introduced in the video after being subjected to double MPEG compression to detect the manipulation. The same authors have performed some modification and suggested a new technique in [133] to check whether a digital video is doubly MPEG compressed or not. Subramanyam et al. [115] have suggested a passive approach for the detection of spatial and temporal copy-paste forgery using video compression artifacts and Histogram of Oriented Gradients (HOG) features. In case of spatial forgery, thresholding algorithm is applied on video frames to divide it into the blocks, after that HOG features were collected from each of the blocks and subsequently matched with other blocks to detect the copy-paste forgery. For temporal forgery, they analyzed the change in GOP structure size and video compression properties. The authors have reported the detection accuracy as a part of a performance measure. The detection accuracy in case of spatial forgery is 96 % for a 60 × 60, and 80 × 80, size forged area and 93.3 % for 40 × 40, size forged area whereas detection accuracy of temporal forgery is 84.5 % for 60 × 60 size forged area and 99 % for 80 × 80 size forged area. Moreover, the same authors have proposed a new approach based on the estimation theory and double compression in [116]. They detected the double quantization and region manipulation in forged video with the help of variation in DCT coefficient and GOP analysis. Labartino et al. [59] have presented a technique to detect and locate the region manipulations forgery in the video using the analysis of Double Quantization (DQ) traces, Histogram of DCT coefficients and Variation of Prediction Footprint (VPF). A method for the detection and localization of insertion/deletion forgery in the videos using double encoding detection is described by Gironi et al. [33]. They used a VPF and DCT coefficients analysis for detecting forgery in the video. Liu et al. [70] proposed a technique based on the sequence of average residual of P-frames (SARP) for the detection of frame deletion forgery in the video. A technique depending on Spatially Constrained Residual Errors (SCREs) of P frames is implemented by Aghamaleki et al. [2] to identify and locate frame insertion/deletion forgery and double compression in a video. The authors investigated the traces of residual error quantization in video frames. The same authors have

Fig. 14 GOP structure

Table 1 Analysis of compression artifacts based forgery detection techniques (QR: Quantization Scale Ratio)

Ref.	Features/Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Wang et al. [130]	Static and temporal artifacts, DCT Coefficients and *P* Frame Prediction Errors	Double MPEG Compression, Frame Deletion & Insertion	MPEG-1 encoded videos	Nil	Sensitive against noise and change in GOP structure. Localization of forgery is missing
Wang et al. [133]	DCT coefficients distribution and GOP analysis	Double MPEG Compression, Region Manipulation, Frame Deletion & Insertion	3 MPEG-2 encoded videos	If QR$<$1.3 then DA=25 %, If 1.3$<$QR$<$1.7 then DA=41.2 %, If QR $>$1.7 then DA=99.4 %	Not suitable if the first compression value is higher than the second compression value. Accuracy depends on the quantization Scale ratio. Beneficial only for the video with a static background & fixed GOP length. Localization of forgery is not precise.
Subrama-nyam et al. [115]	Video Compression Properties and HOG features	Spatial & temporal Copy-Paste Forgery	15 MPEG-2 encoded videos	Spatial Forgery: DA=94.65 % & Temporal Forgery: DA =91.75 %	Forgery localization is not done. Suitable for videos captured through the static camera & with fixed GOP length only. Accuracy decreased when a forged region is small.
Subrama-nyam et al. [116]	DCT coefficients and GOP analysis	Double Quantization & Region Manipulation	Video Trace Library (VTL) [126] (MPEG-2 encoded)	If 1.2$<$QR$<$1.3 then DA=80 %, If 1.3$<$QR$<$1.7 then DA=87 %, If QR $>$1.7 then DA=96 %	Localization is not done. Accuracy decreased when a forged region is small, and it depends on the quantization ratio. Not suitable for the videos having a moving background & variable-length GOP.
Labartino et al. [59]	DQ Traces, DCT Coefficients & VPF	Region Manipulation	Videos from [138] (MPEG-2 encoded)	ROC and AUC	GOP estimation is not feasible in the presence of B-frames. Only work with MPEG-2 VBR coded & fixed-size GOP videos.
Gironi et al. [33]	VPF & DCT coefficients	Frame Insertion & Deletion	14 videos from VTL [126] (H.264, MPEG-4 & MPEG-2 encoded)	TP, TN, & DA	Forgery localization is not precise. Work with fixed size GOP videos. Failed when someone inserts or remove the whole GOP.

Table 1 (continued)

Ref.	Features/Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Liu et al. [70]	P-frame residual error & GOP analysis	Frame Deletion	20 Videos encoded by H.264 codec from VTL [126]	TPR=91.82 %, FPR=5 % & DA=92.08 %	Handle a single type of forgery only. Work with fixed size GOP videos. Localization is not done.
Aghamaleki et al. [2]	SCREs of P frames. DCT coefficient and Traces of Quantization Error	Double Compression, Frame Insertion & Deletion	22 YUV videos from VTL [126] (MPEG-2 encoded)	TPR, FPR, DA=92.73 %	Not suitable for the videos with a moving background. Performance is affected for video having a low compression ratio.
Aghamaleki et al. [3]	DCT Coefficient & Residual Errors	Double Compression, Frame Insertion & Deletion	22 YUV videos from VTL [126] (MPEG-2 encoded)	DA=83.39 % PR=88.4 % RR=90.5 %	Not adequate for the videos with a moving background.
Fadl et al. [29]	Residual Frames & Entropy of DCT Coefficients	Frame Duplication	SULFA [86] & VIRAT Video dataset [82]	On SULFA [86]: RR=98.3 % TPR=99 % F1=98.6 % & On VIRAT [82]: RR=97.1 % TPR=98.2 % F1-Score=97.6 %	Detect the single type of forgery only. Not suited for the video with moving background & variable-length GOP.

also introduced another technique in [3], which consists of three modules, such as detection of double compression, detection of malicious manipulation, and fusion of decisions. In the detection of double compression, The DCT coefficients of I-frames are used as features which are then supplied to the Support Vector Machine (SVM) classifier to classify the single or double compressed video. Whereas, malicious tampering detection module analyzed the time-domain analysis of quantization effects on residual P frame errors to determine the frame insertion or deletion forgeries. Lastly, the output of both the module is fed to the decision fusion module to classify the videos into three types as Single Compressed Videos; Double Compressed Videos with forgeries and Double Compressed Videos without forgeries. The benefit of the both of the proposed technique [2, 3] is that it can work for the video with distinct GOP lengths and structure; however, the performance is affected for the video with moving camera and low compression ratio videos. Fadl et al. [29] have developed an approach based on the concept of residual frames for the identification and localization of digital video inter-frame duplication. The entropy of DCT coefficients in the standard deviation value of each residual frame is calculated, and the similarity among the pairs of feature vectors is explored to detect and locate the frame duplication forgery.

3.2.2 Noise artifacts based techniques

Noise is an essential feature or a clue in the video forensics for the identification of various forgery in the video. Noise artifacts based techniques take the help of sensor artifacts produced by the digital camera. Digital Video Camera usually leaves a characteristic fingerprint in the form of noise which can be used by the researcher to expose the forgery in the video due to that reason someone may also be called it as a camera-based detection technique. The noise artifacts used in video forgery detection is as shown in Fig. 15. Several noises such as Photon Shot Noise (PSN), Fixed Pattern Noise (FPN), Sensor Pattern Noise (SPN), Quantization Noise (QN) and Photo Response Non-Uniformity Noise (PRNU) are used for the detection of forgery in the video. Table 2 shows the analysis of video forgery detection techniques based on noise artifacts.

Mondaini et al. [76] used FPN, PRNU and Self-Building Reference Pattern (SBRP) to identify forgeries such as object insertion, copy-move and frame insertion in a video. The noise is extracted from the video frame, and then the several correlations among them are computed to detect the forgery. The technique is tested on both compressed and uncompressed video, but it works efficiently only for uncompressed video with a stationary background. Hsu et al. [41] used the noise residue correlation at the block level to

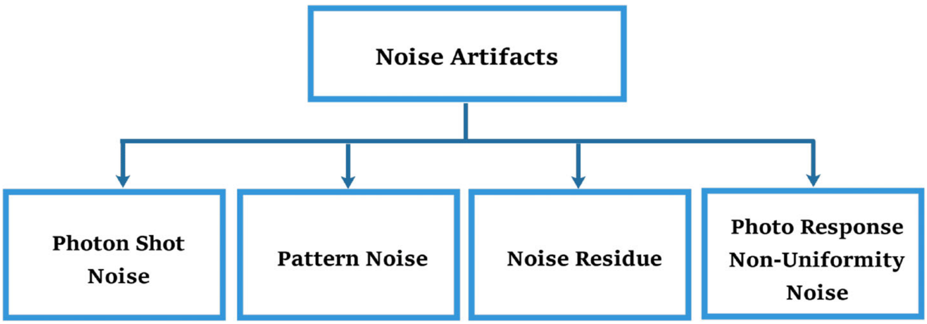

Fig. 15 Noise artifacts for video forgery detection

Table 2 Analysis of video forgery detection techniques based on noise artifacts

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Mondaini et al. [76]	PRNU & FPN	Object Insertion, Copy-Move & Frame Insertion	Own dataset with MPEG compressed videos	Nil	Adequate for videos having a static background & fixed GOP length only. Post-processing operation and MPEG compression affect performance. Reliable only on uncompressed videos. Localization is not done.
Hsu et al. [41]	Noise Residue, GMM, Bayesian Classifier and EM Algorithm	TCP & ETS Inpainting	Personal dataset with MPEG-2 coded video	For TCP: RR=57.76 % PR=96.61 % MR=44.23 % FPR=3.38 % For ETS: RR=74.58 % PR=92.80 % MR=25.41 % FPR=7.195 %	Noise residue extraction is complicated. Sensitive against illumination & quantization noise. Appropriate only for videos having static background & fixed GOP length.
Kobayashi et al. [53]	PSN	Region Manipulation	Created own videos compressed by the lossless huffyuv codec	Recall & Precision	Useful for videos having a static background. Reliable only on videos compressed by lossless huffyuv codec. Not use the spatial relation of pixels. Brightness in the pixel cause degradation of detection.
Kobayashi et al. [54]	PSN & NLF	Region Manipulation	Created own videos compressed by the lossless huffyuv MPEG-2, H.264, Cinepak codec	For huffyuv : TP=0.97 TN=0.98 For MPEG-2:TP=0.46 TN=0.55, For H.264 TP=0.39 TN=0.53, For Cinepack: TP=0.062 TN=0.91	Not suitable for moving scenes or objects. Accuracy depends on the use of video codec. Post-processing operations such as compression, brightness and contrast in the pixel affect the performance.

Table 2 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
chetty et al. [21]	Noise & Quantization Residue Features	Copy-Move	Videos collected from Internet streamed movies	DA=92 %	Suitable only for the video with a static background. Work on a single type of forgery only.
Hyun et al. [45]	SPN & MACE	Upscale-Crop Forgery	Own dataset with 480 MPEG-4 coded videos	TPR=79.86 % & FPR=0.45 %	Not suitable for videos with moving background or objects. Provides an average performance and work well with MPEG-4 encoded videos only.
Ravi et al. [87]	Frame compression noise extracted using HMRF	Double Compression, Frame Deletion & Copy-Move	Video from[138] (MPEG-2 & 4 encoded)	DA=95 %	Performance is depending on the quantization scale. Localization is not done.
Pandey et al. [83]-a	Residual Errors & DCT	Frame Duplication	Own dataset & SULFA [86] (H.264 & MJPEG encoded)	DA=98 % PR=92.45 % RR=99.40 % FP=0.01 %	Accuracy reduces with the increase in compression. Not robust against postprocessing operations. Not suitable for video having a moving background & variable-length GOP structure.
Hu et al. [42]	Camera Noise, Extrinsic Camera Parameters	Region Manipulations	Videos collected from YouTube	Translation Difference (DT) Rotational difference (RD)	Camera parameters affect performance. Handle a single type of forgery only. Provide insufficient validation
Singh et al. [102]	SPN, Local noise variation & Pixel-correlation examination	Upscale-crop & Splicing	Videos from SULFA [86] & VTL [126] (MPEG-2 & 4 encoded)	DA=98 %	Not beneficial for moving background & variable-length GOP structure videos.

Table 2 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Singh et al. [101]	SPNC, CFA & H-DC	Copy- Move	Videos from SULFA [86], VTL [126] (MJPEG & H.264/AVC encoded)	For SPNC: DA=89.9 % to 98.7 % & For CFA-V: DA=83.2 % to 93.3 % & For HD-C: DA=79.1 % to 90.1 %	Handle a single type of forgery only.
Fayyaz et al. [31]	SPN & Noise Residue Correlation	TCP Inpainting	Videos from [41]	DA=96.61 % FPR=3.38 %	Detect a single type of forgery. Not suitable for moving background videos.

locate inpainting forgery like TCP and ETS in a video. The authors worked on the principle that when some frames have tampered, then the correlation values between temporal noise residues changes. Firstly, the video is divided into a series of frames. After that noise residue is extracted from each of these frames, then, these video frames are further partitioned into non-overlapping blocks, and the correlations between every two consecutive frames are calculated. Finally, the forgery present in a video is located by analyzing the correlation of block-level noise residue using Gaussian Mixture Model (GMM) model and the Bayesian classifier. A method based on the noise inconsistencies is introduced by Kobayashi et al. [53] for the identification of forged regions in the video. The photon shot noise is exploited as a piece of evidence, and the linear Noise Level Function (NLF) is formulated to analyze the relationship among the extracted noise to detect the forgery. The same authors have extended the existing work and suggested another method in [54] based on the Nonlinear NLF and inconsistencies in noise to detect the manipulations. The characteristic of photon shot noise is exploited, and correlations among both variance and mean are calculated with the help of nonlinear NLF to expose the manipulations. The framework to handle the copy-move tampering in the video is presented by Chetty et al. [21]. The noise and quantization residue features are obtained from the sub-block of each video frames. These features are then converted into cross-modal subspace to detect the forgery. The SPN based representation method is proposed by Hyun et al. [45] with the help of Minimum Average Correlation Energy (MACE) filter to detect the forged region in the video. This method is also used for source camera identification. In the first stage, the source camera for a given video is identified. Then in the second stage, several forgeries such as partial manipulation, video alternation and upscale-crop are identified by computing the scalar factor and correlation coefficient. Video forgery detection technique is proposed by Ravi et al. [87] for frame deletion and copy-move forgery by identifying double compression. The compression noise is used as a feature which is extracted from the video frames by the modified Huber Markov Random Field (HMRF) prior model. The extracted noise can be modelled as a first-order Markov features which are then given to the SVM classifier to detect the forgery. Pandey et al. [83]-a have designed an approach for the detection of temporal copy-move forgery (i.e., frame duplication) in the video using wavelet denoising and noise residue-based techniques. Hu et al. [42] have developed a technique to detect the region tampering in digital video using the properties of extrinsic camera parameters. Firstly, each of the video frames is divided into several block areas, followed by the calculation of extrinsic parameters from each of these blocks. Then differences among these parameters are computed. Finally, a certain threshold is chosen to detect the manipulations. Singh et al. [102] have proposed the techniques to detect intra-frame forgeries such as upscale-crop (outer parts of the frames are cropped out) and splicing forgery using pixel-correlation examination and noise-inconsistency investigation. For that, they used the resampling detector, which is referred to as Modified-Gallagher (MG) Detector and F-MG Detector (Fractional MG). In addition to this, authors have presented three schemes in [101] to detect and localize the copy-paste forgery in digital video. In the first scheme, Sensor Pattern Noise Correlation (SPNC) is used to detect and locate the manipulation. In the second, Color Filter Array Artifacts (CFA-V) is used to expose the manipulations in uncompressed frames. The final scheme is a Duplicate Cluster Detection Scheme (H-DC) based on the concept of Hausdorff distance-based pixel-clustering to identify the manipulation. The presented technique able to detect the forgery from MPEG-2, 4, MJPEG and H.264/AVC encoded videos, captured with static and moving cameras and it is independent of GOP structure length. With the use of SPN and noise residue correlation, Fayyaz et al. [31] developed the technique to detect the temporal copy-paste inpainting forgery. The noise residue patterns are extracted from each of the video frames and then compared it with the collected SPN using adaptive DCT filtering to detect the forgery.

3.2.3 Motion features based techniques

Motion-based features are time-dependent features in the digital video which define the relationships among the adjacent frames. When forgery is performed in the digital video, then motion features and relation among the adjacent frames are going to be changed, this used as a clue to identify the forgery in the video. Motion features for video forgery detection is shown in Fig. 16. The motion-based features are captured in the form of Motion Residual, Optical Flow Coefficients, Motion Vector Pyramid (MVP), and Motion Compensated Edge Artifacts (MCEA). The MCEAs are special artifacts that occur in videos that are compressed using block-based motion-compensated frame prediction coding algorithms. Successive video frames are decoded with the aid of previously decoded frames during motion-compensated frame estimation, which allows successive video frames to become dependent on each other. Inter-frame forgeries break these associations or comparisons, resulting in even more visibility in the current block boundary objects in the video frames. The spike in block boundary objects, known as MCEA, will help detect inter-frame forgeries. Another useful forensic aspect that enables the detection of inter-frame forgeries is optical flow, that refers to the pattern of the apparent movement of objects, edges, and surface within successive video frames. In a genuine video, optical flow differences between successive frames appear more or less constant, in case if some inter-frame manipulation is performed on video, the optical flow starts to show such anomalies that can act as the fingerprint. Velocity field relates to the disturbance between neighbouring video frames induced by time separation. The velocity field tends to follow a consistent pattern in a genuine video, whereas it gets disturbed in case of forgery is done on the video. The analysis of video forgery detection techniques based on motion features is shown in Table 3.

Wang et al. [131] have suggested an adaptive motion algorithm to identify region manipulation forgery in de-interlaced and interlaced video. They analyzed the changes in correlation introduced by the de-interlacing algorithm to identify the forgery in the de-interlaced video. Whereas to identify the forgery in interlaced video, they measured the interfiled and inter-frame motion. MCEA based technique is presented by Su et al. [114] for frame deletion forgery in digital video. They explained MCEA error which is produced after frame deletion manipulation in the video due to the effect of a decrease in temporal

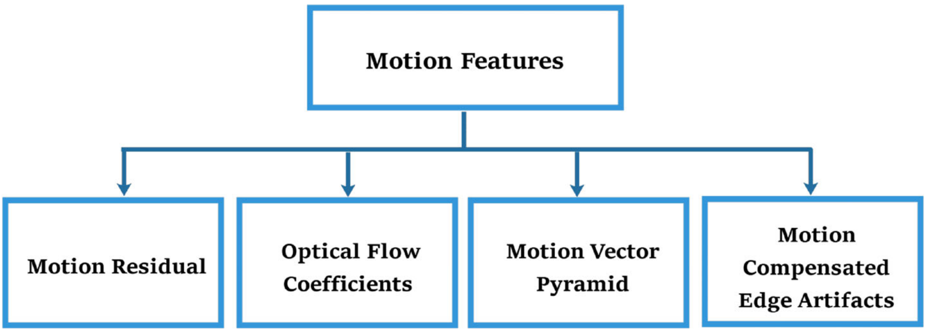

Fig. 16 Motion features for video forgery detection

Table 3 Analysis of video forgery detection techniques based on motion features (LA: Localization Accuracy)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Wang et al. [131]	Inter-Field & Inter-frame Motions. Correlation Factor	Region Manipulation	Own dataset	For de-interlaced un-compressed videos: DA=100 % & For MPEG compressed videos DA are 97 %, 96.1 %, and 93.3 % at bit rate 9, 6, and 3 Mbps, respectively.	Sensitive against the compression and noise. Not suitable for the videos with moving background & variable-length GOP structure.
Su et al. [114]	MCEA	Frame Deletion	5 MPEG-2 encoded videos	Nil	Not Suitable for videos with variable length GOP structure & slow-motion videos. Failed if entire GOP is deleted.
Dong et al. [28]	MCEA & FFT	Frame Insertion & Deletion	4 videos from VTL [126] encoded with MPEG-2 codec	Nil	Appropriate for videos with static background only. Change in GOP structure affects the performance. Failed if the frames deletion count in a video is an integer multiple of GOP. Not suitable with the H.264/ AVC encoded videos.
Kancherla et al. [49]	Motion residue, Markov models & SVM	TCP & ETS Inpainting	20 videos from [80, 81], [98, 124]	DA=87 % PR=89 % RR=86.5 % ROC curve AUC=0.9479	An experiment is performed on a small dataset. Feature extraction is not done at the bit-stream level. Not appropriate for the videos having a moving background & variable-length GOP structure.

Table 3 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Li et al. [61]	Block-Based Motion Estimation	Object Removal Forgeries	Videos taken from Internet	Nil	Only focused on the detection of rigid object removal. Work for the videos with a stationary background & fixed-length GOP structure. Improvement is needed for compressed videos.
Bestagini et al. [13]	Analysis of the Footprints left on the Residual	Adding & Removing the objects	Created own dataset (Named as REWIND) (videos with H.264/ AVC codec)	TP=0.62 FN=0.38 TN =0.94 FP=0.06 ROC curve AUC=0.91	Performance depends on the quantization parameter, and it is affected when quantization is accepted in the second step of encoding. Not suitable for videos with moving background & variable-length GOP structure.
Chao et al. [17]	Optical Flow Coefficients	Frame Deletion & Insertion	Videos from KTH [58] & TRECVID [121]	For Insertion: PR=98 % RR=95 % & For Deletion: PR=89 % RR=85 %	Need improvement in case of frame deletion forgery. Not ideal for the video having a moving background & variable-length GOP structure.
Wang et al. [134]	Optical Flow Coefficients	Frame Deletion, Insertion & Duplication	40 Videos from TRECVID [121] (MPEG-2 codec) & SULFA [86] (H.264 & MJPEG encoded videos)	For Insertion: DA=85 % LA=100 %, For Deletion: DA=72 % LA=96.9 % & For Duplication DA=82.5 % LA=86.2 %	DA is lower for frame deletion. In contrast, LA is lower for frame duplication. Improvement is needed in optical estimation method. Not work for videos with moving background & variable-length GOP structure.

Table 3 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Feng et al. [32]	Distinctive Fluctuation Feature of Motion Residual	Frame Deletion	VTL [126] (H.264/ MPEG-4 codec)	TPR=90 % FAR<=0.8 %	Not robust against the sudden lightening change and zooming. Accuracy decreases with the increase in bit rate. Not appropriate for videos with moving background & can handle a single type of forgery.
Wu et al. [137]	Block-based Cross-correlation, Velocity Field Sequence and ESD test & local noise variation	Frame Deletion & Duplication	4 videos from TRECVID [121] (MPEG-2 encoded)	For Deletion: DA=85 %, For Duplication: DA=80 % & Video is tampered or not, DA=96.3 % FP=10 %	Accuracy decrease with the increase in compression. Performance depends on the quantization scale. Not adequate for videos with a moving background.
Wang et al. [129]	Optical flow coefficients	Frame Insertion & Deletion	598 video divides in five groups	DA for separate forgery in X & Y direction= 94.01 % & 92.93 % Resp., DA for mixed forgery in X & Y direction= 90.77 % & 91.31 % Resp.	Localization is missing. Not adequate for the videos with moving background & variable-length GOP structure. Computationally inefficient as it required to calculate the optical flow consistency features in both X direction & Y direction.
Tan et al. [119]	Motion Residual feature, GOP Structure analysis and Two ensemble classifier	Objecet Insertion & Deletion	SYSU-OBJFORG [20]	DA=80 %	Need to improve the DA. Not appropriate for the videos with moving background & distinct length GOP structure. Localization is missing.
Bidokhti et al. [14]	Optical Flow Coefficients	Copy-Move & Frame Duplication	REWIND [88] & SULFA [86] dataset	DA=90.67 %	Highly sensitive against the Region of Interest (ROI) selection. Not suited for variable-length GOP's with high motion. Failed if the area of forged regions is a multiple of GOP's length.

Table 3 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Zhang et al. [150]	MVP & VF Consistency	Frame Deletion & Duplication	30 videos (MPEG-2 encoded) from TRECVID[121]	Deletion: DA=90.67 % LA=96.66 % & Duplication: DA=83.33 % LA=100 %	Appropriate for MPEG-2 encoded videos with static-background & fixed-size GOP only. Improvement is needed in DA, especially in case of duplication. Not robust against videos re-encoded with different coding standards.
Yu et al. [146]	Analyzing Abrupt Changes in Video Streams	Frame Deletion	VTL [126] (encoded with H.264 codec)	PR, RR, F1-Score	Failed if the count of deleted frames is small. Failed when a forged video is of slow motion & having a moving background.
Chen et al. [20]	Motion Residual feature and Three ensemble classifier	Object Removal & Insertion	SYSU-OBJFORG [20] (Videos encoded with H.264/MPEG-4 codec)	PFACC, FFACC, DFACC, FACC, VACC, Precision, Recall and F1-Score	Exact localization is not done. Not adequate for high resolution & high bit rate videos. Not suited for the videos collected from a moving camera. Work only with a fixed size GOP videos.
Singh et al. [103]	Optical Flow Coefficients & Prediction Residual Examination	Frame Insertion, Deletion & Replication	10 videos encoded with H.264/AVC and MJPEG codec. Used SULFA [86] & VTL [126]	For Insertion: DA=98.2 %, For Deletion: DA=98.6 %, For Replication: DA=98.3 %	Not suitable for the videos with a moving background. Performance affected in the case of multiple compressed videos.

Table 3 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Kingra et al. [52]	Motion & brightness gradient features, Optical Flow Coefficients and Prediction Residual Analysis	Frame Insertion, Deletion & Duplication	DIC-Panjab University (videos encoded with H.264 & MPEG-2 codec)	Avg. DA=83 %, Avg. LA=80 %, Avg. TPR=81 %, Avg. FPR=6 %	Not suitable for the videos having a high illumination, moving background & variable-length GOP structure.
Sitara et al. [107]	Inconsistency in Velocity Field and VPF	Frame Insertion, Deletion, Duplication & Shuffling	78 videos from VTL [126] (encoded with MPEG-4 & H.264 codec)	PR=98.3 %, RR=92.3 %, F1-Score=95.2 %, DA=92.3 %	Testing is not done on videos with different Quantization Scale. Sensitive against sudden zooming. Not suitable for the videos with a moving background.
Pu et al. [85]	GSSIM & Optical flow	Frame Deletion	400 videos from SULFA [86] & CDNET [135]	On SULFA: PR=99.95 %, RR=99.45 % & ON CDNET: PR=94.62 %, RR=92.11 %	Work on a single type of forgery only. Not suitable for night videos. Work with a static background & fixed-length GOP video.

correlation. One more MCEA based technique is designed by Dong et al. [28] to detect the inter-frame video forgery like frame insertion/deletion. The MCEA value of each P frame in the video is extracted, and the Fast Fourier Transform (FFT) is used onto the difference of MCEA values between adjacent P frames. Then check for the presence of spikes in Fourier transform (if present then a video is tempered else it is authenticated). The inpainting forgery such as TCP and ETS in the videos are detected by Kancherla et al. [49] using a Markov model on extracted motion-based features in the videos. The salient motion-based features in a video using motion extractor and Markov model is extracted. The SVM algorithm is then used to obtain a binary classification on these extracted features. The Block-based motion estimation algorithm is presented by Li et al. [61]. The authors have detected the object removal forgeries in digital videos. They have analyzed the fact that if the certain object is deleted from the video, then the motion vector is changed. The motion information in the form of the motion vector is extracted as a clue of tampering from the adjacent video sequences to detect the forgery. Then, the original region is differentiated from the manipulated region using the orientation and magnitude of the motion vectors. Based on the analysis of the footprints left on the residual, Bestagini et al. [13] have proposed an algorithm that detects the tampering such as adding or removing certain objects from videos. They also made an enlargement of the SULFA database [86] by adding more forged videos in it. The authors have reported some parametric value such as TP, FN, TN and FP which are 0.75, 0.25, 0.97 and 0.03 respectively for the video which is not recompressed, 0.71, 0.29, 0.98 and 0.02 respectively for the video with Quantization Parameters QP =10, 0.58, 0.42, 0.96 and 0.04 respectively for the video with QP =15 and 0.44, 0.56, 0.84 and 0.16 respectively for the video with QP =20. Chao et al. [17] have presented an inter-frame forgery (insertion and deletion forgery) detection method for the digital video using an optical flow consistency algorithm. The window-based rough detection model is designed for the insertion forgery. Whereas, the frame-to-frame mechanism and double adaptive threshold-based detection model is designed for detecting the frame deletion forgery. Wang et al. [134] have also developed an optical flow-based algorithm for forgery detection and localization in digital videos by analyzing the discontinuity points and optical flow sequence. They extracted optical flow variation sequence from adjacent frames to locate discontinuity points and detected the forgery such as a frame insertion, deletion, and duplication. The algorithm for handling the frame deletion forgery in the video is proposed by Feng et al. [32] based on the total motion residual. They exploited the distinctive fluctuating feature of motion residual to detect the deletion forgery and used the adaptive threshold method to locate it. The testing is performed on CBR and VBR encoded videos with both fixed and variable-length GOP structure are taken from VTL [126]. Wu et al. [137] have developed an algorithm to detect forgeries such as frame deletion and duplication in the digital video. They used block-based cross-correlation on the video to find a velocity field sequence. The generalized Extreme Studentized Deviate (ESD) test is used to detect and locate the forgery present in the video. An Inter-frame forgery detection method for digital video is created by Wang et al. [129] using an optical flow consistency. The optical flow values between each of the adjacent video frame in both x and y direction are calculated. The computed values are then given to the SVM to differentiate the forged and original video. The authors reported the classification accuracy for the single type of forgery in the x-direction for 25 frame insertions, 100 frame insertions, 25 frame deletions, and 100 frame deletions are 98.41 %, 98.20 %, 86.82 %, and 92.61 % respectively. Whereas, the classification accuracy for the single type of forgery in y-direction for 25 frame insertions, 100 frame insertions, 25 frame deletions, and 100 frame deletions are 98.60 %, 98.54 %, 86.02 %, and 88.56 % respectively. For the two types of forgery, the classification accuracy of 25 frame insertion and deletion in both x and y direction are 91.72 % and 90 % respectively whereas for 100 frame insertion and deletion in both x and y direction are 89.83 % and 92.63 % respectively. The technique based on GOP structure for object-based manipulation (adding or erasing moving object) detection in a digital video is proposed by Tan et al. [119]. They created frame manipulation detector with the use of motion residual extracted from video frames. Then CC-PEV feature set is utilized to obtain the feature vector from each of the motion residual. Later, these feature vectors are given to two ensemble classifier which categorized the video into pristine, double compressed or forged one. Based on the Lucas Kanade optical flow Bidokhti et al. [14] have developed a technique to expose the copy-move and frame duplication forgery in the video. Firstly, the video frames are separated into two parts. After that, an optical flow coefficient is calculated between these video frames. Finally, the forgery in a video is identified if any unusual changes are observed in optical flow coefficients. The MVP (Motion Vector Pyramid) consistency and it's Variation Factor (VF) is used by Zhang et al. [150] to detect and locate the frame deletion and frame duplication forgery in the video. They used discontinuity points in the VF sequence as a clue for detecting a forgery in the digital video. The method divided into two stages 1) Features extraction 2) Discontinuity point detection. The MVP sequence with it's associated VF is computed for the subsequent frames in a video in the first stage. Moreover, in the subsequent stage, forgery is detected and localized with the use of a modified generalized ESD test. Yu et al. [146] have proposed the approach for the identification of frame deletion forgery in the video by analyzing abrupt changes in video streams. The authors have used two features to find out the magnitude difference in prediction residual (PR) and the Number of Intra Macroblocks (NIMB's). Based on these features, the fused index is constructed to detect frame deletion forgery. The passive forgery detection algorithm is developed by Chen et al. [20] to identify and localize the object-based tampering (Insertion or removal of objects) in the video using motion residual features. The frame manipulation detector is used to find out the residual motion feature left in video frames produced by the unethical operations. Then, SPAM, CC-PEV, CDF, SRM, CF*, J + SRM, and CC-JRM feature sets are used to create the feature vector which is obtained from each of the motion residual.[1] Then the ternary classifier (Ensemble Classifier) is used which takes these feature vectors as input and categorizes the corresponding video into a pristine, double compressed, or forged one. Singh et al. [103] developed the forensic system based on optical flow and the prediction residual to handle the frame insertion, deletion, and replication forgery in the video. The optical flow analysis-based technique is used here for frame insertion and deletion detection, where they focused on the brightness gradient component of optical flow. However, the prediction residual examination scheme is used to detect and localize the replicated frames. The forgery detection technique using the optical flow gradients features and the analysis of prediction residual is presented by Kingra et al. [52]. The technique can identify and locate the frame deletion, insertion and duplication in the videos. They evaluate the fact that the temporal correlations among the adjoining frames are disrupted when the video is manipulated. The window-based concept is used to locate the forgery. The proposed scheme is specifically designed for H.264 video and MPEG-2 codec. It works well for both slow and fast motion video, while the detection performance is slightly affected when the video is subject to high illumination. Sitara et al. [107] have developed a technique to expose the frame deletion, insertion, duplication, and shuffling

[1] The abbreviations for the feature discussed earlier mentioned in http://dde.binghamton.edu/download/feature_extractors/.

forgeries in the videos using inconsistencies in the velocity field and VPF. The Generalized Extreme Studentized deviate (ESD) algorithm is designed by the authors to locate the forged places in the video. The technique is capable of identifying forgery even if the complete GOP's Structures are deleted and also for the adaptive GOP structure. The approach based on spatial constraints and stable feature to expose the frame deletion forgery in the video is proposed by Pu et al. [85]. Initially, they obtained a Quantitative Correlation Rich Region (QCRI), then optical flow information is calculated to identify suspicious forged points. At last Gradient Structure Similarity Feature (GSSIM) are calculated to finalize the forgery. The proposed approach is independent on the frames deletion count, and it is robust against the attacks like noise, filtering and blur.

3.2.4 Statistical features based techniques

Statistical feature-based or pixel-based techniques for the video forgery detection look at statistical attributes/properties of objects, pixel-level variance and correlations among frames. This technique is also called Geometric/physics inconsistencies-based techniques as it deals with the inconsistencies (such as lighting, brightness, shadows, *etc.*) in the video frames. The statistical attributes may be changed after performing the forgery in the video, which is then investigated to detect the manipulations. Figure 17 shows the statistical features used in video forgery detection. Table 4 shows the analysis of video forgery detection techniques based on statistical features.

Based on the temporal and spatial correlations, Wang et al. [132] have exploited the correlation coefficient as a measure to detect the forgery in the video. Based on ghost shadow artefact, Zhang et al. [148] have presented a technique to identifies the video inpainting forgeries such as TCP and ETS. The statistical properties of the object based on the Adjustable Width Object Boundary (AWOB) algorithm is used by Chen et al. [19] to identify the object insertion or removal forgery in the video. The contourlet coefficient and

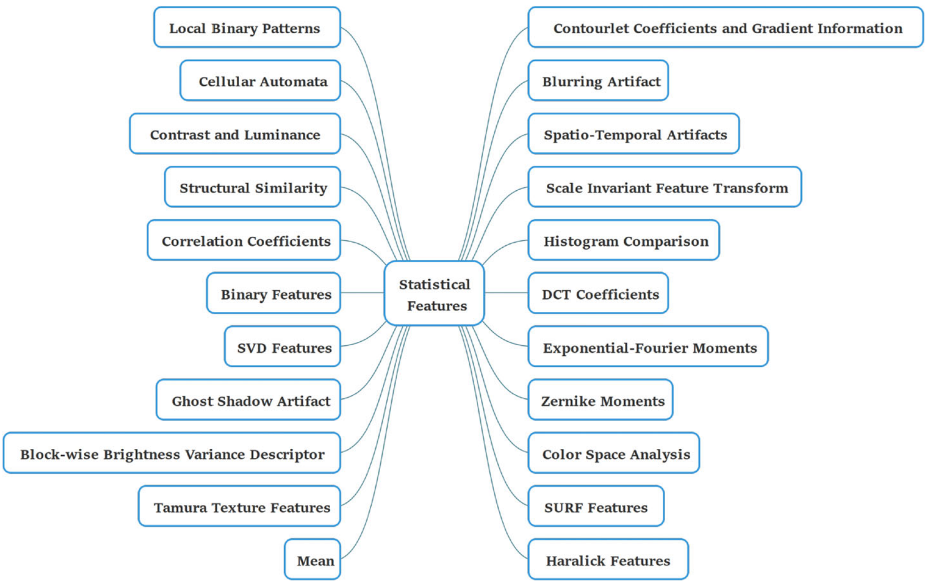

Fig. 17 Statistical features used in video forgery detection

Table 4 Analysis of video forgery detection techniques based on statistical features

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Wang et al. [132]	Temporal & Spatial Correlations	Frame Duplication & Copy-Move	Videos Recorded by SONY-HDR-HC3 camera	DA	Accuracy is affected by a change in bit rate and region size. Not suitable for the videos with moving background & variable-length GOP structure. Exact Localization is missing.
Zhang et al. [148]	Ghost Shadow Artifact	TCP & ETS Inpainting	10 MPEG-2 encoded videos from [41]	Nil	Only suitable for videos with a stationary background & fixed-length GOP structure. Precise localization of forgery is missing.
Chen et al. [19]	AWOG, Contourlet Coefficients & Gradient Information	Object Insertion or Removal	9 AVI & WMV format videos	DA=96.83 % PR=96.28 % RR=96.43 % FPR=1.18 %	Work for the video with statics background only. Used features depend on training samples. Localization is missing.
Hu et al. [43]	TIRI-DCT	Frame Duplication	Personal dataset with 3 CIF format video clips	TPR=100 % FPR=0 %	Works for the videos with a stationary background & fixed-length GOP. Localization is missing.
Lin et al. [69]	Histogram difference between adjacent frame.	Frame Duplication	Personal dataset with 15 video clips	PR=84.9 % RR=100 % DA=100 %	Need to combine other features to improve the efficency. Provides an average performance. Not suitable for videos with moving background & variable-length GOP structure.
Liao et al. [66]	Tamura Texture Feature	Frame Duplication	Personal dataset with 10 videos	PR=99.6 % RR=100 %	Need to combine other features to reduce the computation time. Detect a single type of forgery only.

Table 4 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Lin et al. [67]	STSA artifacts	TCP & ETS Inpainting	1 video	Nil	Moving object forgery and multiple objects removal is not handled. Not suitable for the videos with moving background & variable-length GOP structure.
Lin et al. [68]	STCA artifacts	TCP & ETS Inpainting	Personal dataset with 18 videos encoded with MJPEG codec	PR=93.6 % RR=80.2 % F1-Score=85.5 %	Performance decreased with the increase in the compression of video. Not useful for the video encoded with a modern codec such as MPEG-2, MPEG-4 and WMV-9. Only suitable for videos with a fixed-length GOP structure.
Li et al. [60]	Structural Similarity	Frame Duplication	15 videos captured from mobile and digital camera.	PR=99.7 % RR=100 %	Not suitable for the videos with a long-time static scene, with a moving background & with a variable length GOP structure. Computational time is high.
Zheng et al. [153]	BBVD	Frame Insertion	240 AVI format Videos from KTH Database [58] & TRECVID [121]	For Detection PR=94.09 % RR=98.67 % & For Localization PR=79.45 % RR=89.23 %	Forgery Localization accuracy is low. The efficiency is decreased if the frame inserted or deleted count is less than 25. Only works for videos with a static background & fixed-length GOP structure.
Wang et al. [128]	CCCoGV	Frame Insertion & Deletion	Created personal data-set (598 videos)	For Insertion DA=99.28 %	Only works for videos with a stationary background & fixed length GOP structure.

Table 4 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Yin et al. [145]	Correlation coefficients and NTF and	Frame Insertion & Deletion	Personal dataset (Video Recorded on DV Sony DCR-HC33E)	For Insertion PR=100 %, RR=99 % & For Deletion PR=88.64 %, RR=88.67 %	Frame deletion performance needs to be improved. Efficiency is affected if the frame inserted or deleted count is less than 25. Not suitable for the video with moving background & variable-length GOP structure.
Chittapur et al. [22]	Mean	Region Manipulation	100 Videos collected from different source	Nil	Only suitable for videos with a static background & fixed-length GOP structure. Handle a single type of forgery. Insufficent validation.
Tralic et al. [120]	CA and LBP	Frame Duplication	SULFA [86]	PR=100 % RR=97.13 % SP=100 %	Not work adequately when multiple frames are duplicated. Suitable only for the videos with a static back-ground & fixed-length GOP structure.
Zhang et al. [151]	LBP and QCCoLBPS	Frame Insertion & Deletion	599 video from KTH Database [58]	For Insertion PR=98.7 % RR=94.91 %. For Deletion PR=91.79 % RR=89.47 %. For Mixed PR=88.16 % RR=85.80 %	Performance affected when the frame insertion or deletion count is less than 25. Not adequate for the video with moving background & variable-length GOP structure. Forgery localization is missing.
Singh et al. [105]	Mean and Frame Resi-due Correlation	Frame Duplication	Own dataset	For Stationary Camera DA=98.1 % & For Moving Camera DA=99.5 %.	Handle a single type of forgery. Appropriate only for videos with a fixed-length GOP structure.

Table 4 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Pandey et al. [83]-b	SIFT feature and KNN matching	Copy-Move	SULFA [86]	DA=99.9 % PR=95.75 % RR=100 % FP=0.001 %	Accuracy affected when a forged region is small. Not robust against any intentional attacks. Not suited for the video with moving background & variable-length GOP structure.
Su et al. [110]	SVD Features	Moving Object Removal	Personal dataset (20 Videos) & downloaded from Internet.	PR=92.2 % RR=90.5 % DA=89.6 %	Detection accuracy is decreased if the deleted object is small and fast-moving. Longer detection time. Suitable only for the videos with a static background & fixed-length GOP structure.
Bagiwa et al. [10]	Correlation of blurring artifact	Chroma Key Forgery	Personal dataset (MPEG-4 encoded 20 video)	TPR=91.12 % FPR=1.95 %	Performance decreased if the background used in the video is green or blue. Only useful for videos with a static background & fixed-length GOP structure.
Xu et al. [139]	Correlation Coefficients & Histogram Intersection	Frame Insertion, Deletion & Duplication	Personal Dataset (Videos recorded by Logitech C270 Digital camera)	For Insertion PR=90.4 %, RR=90.4 %, For Duplication PR=94.4 % RR=88.2 % & For Deletion PR=95.2 % RR=82.6 %	Not suitable when the frame count is less than 8. Performance affected when compression time is more than three. Work for the videos with a static background & fixed-length GOP only. Forgery localization is missing.
Li et al. [65]	Consistency of QoMSSIM	Frame Insertion & Deletion	598 video from KTH Database [58]	For Frame Insertion: Classification Accuracy=98.79 %, For Frame deletion: Classification Accuracy=92.83 %	Performance affected when frame insertion or deletion count is less than 25. Only suitable for videos with a static background & fixed-length GOP structure. Localization is missing.

Table 4 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Mathai et al. [74]	Statistical Moments Features and Normalized Cross-Correlation	Copy-Move	SULFA dataset [86]	For Detection: TP=0.718 FN=0.28 TN=0.768 FP=0.23 & For Localization: TP=0.81 FP=0.11	Detection accuracy is based on the window size. Need to improve the localization accuracy. Not appropriate for the videos with moving background & variable-length GOP structure.
Yang et al. [140]	SVD Features	Frame Duplication	Videos from SULFA [86], Movie Scenes, CCTV and fixed camera.	DA=99.10 % PR=98.20 % RR=100 %	Not worked when duplication performed in a different order, and frame duplication count is smaller than the window size. Not suited for the videos with moving background & variable-length GOP structure. Detect a single type of forgery.
Liu et al. [71]	ZOCM	Frame Duplication, Insertion & Deletion	60 videos from SULFA dataset [86]	PR=97.5 % RR=99.2 %	Failed for the videos with moving background & different GOP length. Localization is missing.
Liu et al. [72]	Luminance & Contrast. 3FAT & GMM	Blue Screen Compositing Forgery	100 videos captured by SONY HDR-XR160E	TP=97.3 % TN=92.2 % FP=2.7 % FN=7.8 % DA=94.75 %	Not capable of locating smaller size forged regions. Not useful for video with a fast-moving background & variable-length GOP structure.
Bozkurt et al. [15]	DCT coefficients, Correlation Analysis (Forgery Line), Hough Transform	Frame Duplication	Videos from SULFA [86] (MPEG-4 encoded)	DA=98.64 % PR=98.12 % RR=97.25 %	Detect a single type of forgery. Not suitable for moving background & different GOP length videos.
Ulutas et al. [122]	Extracted binary features	Frame Duplication & Frame mirroring	Own dataset (Videos from SULFA [86])	DA=99.35 % PR=99.98 % RR=99.30 %	Suitable for the videos with a stationary background & fixed GOP length only.

Table 4 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Ulutas et al. [123]	SIFT and BoW features	Frame Duplication	Own dataset (Videos from SULFA [86])	DA=97.54 % PR=98.37 % RR=98.38 %	Only handle a single type of forgery.
D'Amiano et al. [24]	Zernike Moments and Patch-match Algorithm	Copy-Move	GRIP Copy-move dataset [35] & RE-WIND dataset [88]	For Basic 2-D F-score=83 %, For Basic 3-D F-score=76 %, For Fast 2-D F-score=79 %, For Fast 3-D F-score=75 %	Localization accuracy is low. Useful for the video with a static background & fixed GOP length only.
Zhao et al. [152]	HSV, SURF features & FLANN matching	Frame Insertion, Deletion & Duplication	Own dataset of 10 video	DA=99.01 % PR=98.07 % RR=100 %	Failed to work with shots, including scene changes. Suited for the video with a static background & fixed GOP length only.
Su et al. [111]	EFMs	Copy-Move	Videos taken from SULFA [86] & Internet	DA=93.1 % TPR=93 % TP=96.9 % TN=89.3 % FP=3.1 % FN=10.7 %	Detect the single type of forgery. Not suitable for the video with moving background & variable GOP length.
Su et al. [109]	MI-SIFT	Copy-Move Forgery with mirror operation	SULFA [86] (MPEG-2 encoded)	DA=92.6 %	Only handle a single type of forgery. Not suitable for video with a dynamic background & variable GOP length.
Su et al. [112]	Energy Factor & AVIBE	Object Removal	SULFA [86] & SYSU-OBJFORG [20]	With static background: DA=93.17 % ; With a complex background: DA=86.58 %	Only handle a single type of forgery. Localization is not precise. Accuracy affected when the selected region is too small.
Wei et al. [136]	Multi-Scale Normalized Mutual Information and Correlation	Frame Duplication, Insertion & Deletion	8 Videos from VTL [126] and 1 self-shoot video	DA=93.33 % PR=96.55 %	Only useful for videos with a static background & fixed GOP length.

Table 4 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Singh et al. [100]	Mean & Correlation	Frame Duplication & Copy-Move	24 videos from SULFA [86] & 6 videos from Internet	Frame duplication: PR=100 %, RR=99 %, DA=99.5 %, F1-Score=99.4 % Copy-move: PR= 100 %, RR=93.3 %, DA= 96.6 %, F1-Score=96.5 %	Algo I is not suitable when frame duplicated count is less. Algo II is unable to detect the forgery when the region is too small. Not useful for variable-length GOP video.
Bakas et al. [12]	Haralick Correlation Inconsistency	Frame Insertion, Duplication and Deletion	17 videos from SULFA [86] and 13 from VTL [126].	PR= 98 %, RR= 97 %, F1-Score=97 %	Performance affected for extremely fast-moving background videos.
Kharat et al. [51]	SIFT features	Frame Duplication	Own dataset of 20 video	For Uncompressed video PR=99.94 %, RR=99.71 %, DA=99.70 % For compressed video PR=100 %, RR=99.71 %, DA=99.76 %	Not suitable for moving the background video. GOP size detail is not mentioned.
Bai et al. [11]	Spatio-temporal LBP	TCP & ETS	7 videos from SULFA [86], 8 videos from Static camera and 5 videos from moving camera	For TCP: PR= 96.14 %, RR= 87.33 %, F1-Score=91.43 % For ETS: PR= 89.99 %, RR= 84.02 %, F1-Score=86.83 %	Shaking and slight rotation affect the performance. Work with a static background & fixed-length GOP video.
Aparicio et al. [8]	Block correlation matrix	Copy-Move and Frame Duplication	10 Video from REWIND [88]	ROC curve	Work with a static background & fixed-length GOP video.

Table 4 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Saddique et al. [94]	DOCFs and CCD-DRLBP	Copy-Move and Splicing	Videos from Hsu [41], Bestagini [13], [86] & [9]	TPR=96.5 %, TNR=93.6 %, DA=96.68 %	Work with a static background & fixed-length GOP video.
Aloraini et al. [5]	Laplacian Pyramid	Object Removal	Videos from [20]	For uncompressed videos: PR=93.3 %, RR=93.3 %, F1-Score= 93.3 % & For Compressed MPEG-4 Videos:PR=92.8 %, RR=92.8 %, F1-Score= 92.8 %	Handle a single type of forgery. Not suitable for moving the background video. GOP size detail is not mentioned.
Aloraini et al. [6]	Mean and Variance , Sequential and Patch analysis,	Object Removal	SYSU-OBJFORG [20]	For Uncompressed video PR=96.65 %, RR=97.78 %, DA=96.70 % For compressed video PR=95.51 %, RR=94.44 %, DA=94.97 %	Not suitable for moving the background video. GOP size detail is not mentioned. Handle a single type of forgery.

gradients information features are extracted from the video frames to identify the manipulations. These extracted features are then supplied to the SVM to distinguish the forged and original objects. Frame duplication detection technique is developed by Hu et al. [43] with the help of video sub-sequence fingerprints. First, the video is divided into the series of frames and formed the Temporally Informative Representative Images (TIRI) of each frame. Then, TIRI is split into the overlapping blocks, and the DCT coefficient is extracted from each of these blocks. Finally, hamming distance is computed to check the similarity among the frames to detect the forgery. They considered TPR and FPR parameters to assess the effectiveness of their method. The average TPR and FPR value without post-processing operations are 100 % and 0 % respectively for the block size 4 and 8 whereas the FPR values are get changed to 55.55 % for the block size 16. The average TPR and FPR values for the videos with a change in brightness and MPEG compressed video are 94.31 %, 0.33 % and 49.31 %, 0.33 % respectively. The new technique is proposed by Lin et al. [69], for the frame duplication detection and localization using spatial and temporal analysis. The technique works in four stages. The first stage is candidate segment selection, where the histogram difference among the adjacent frames in Red Green and Blue (RGB) color space is used as a forensic feature. The second stage is spatial similarity analysis, where the high correlation between the two frames is observed using the block-based algorithm. The third stage is to create a classifier for detecting the duplication forgery, and the last stage is to perform post-processing. Liao et al. [66] have proposed a technique for identifying and locating the frame duplication forgery in the digital video with the use of Tamura Texture Features (TTF). Firstly, TTF features (like contrast, directionality, and roughness) are extracted from each of the video frames to generate an eigenvector matrix. After that, the dictionary ordering concept is applied to sort these eigenvectors to calculate the variation between the eigenvector and their neighbour vectors. Finally, the difference among these eigenvectors is observed to check the duplication forgery. The Spatio-temporal slices are extracted and analyzed by Lin et al. [67] to identify and localizes the inpainting forgeries such as TCP and ETS. The approach is divided into two parts: Spatio-Temporal Artifact Analysis (STSA) and Refinement. The STCA from the video frames is extracted, and abnormal regions with high inconsistency or similarity are analyzed. Then, the map of the Whole Spatio-Temporal Slice Artifacts (WSTSA) is obtained. At last, the refinement process is applied with the use of the WSTA map to match every Spatio-temporal slice artifact to detect the forgery. The limitation of their approach is that it is not suitable for multiple object removal forgery. To overcome the flaws, the same authors in [68] modified the existing approach to identify and localizes the inpainting forgeries such as TCP and ETS in the video. They filled the area left after the object removal forgery, and design a new approach depends on coherence examination to handle the manipulated areas in digital video. The technique has experimented on a set of 18 test videos.[2] Although it detects the multiple object removal forgery, the performance of their technique is affected by an increase in the compression of video. Based on structural similarity Li et al. [60] have suggested a method to detect and locate the frame duplication (alternatively called as temporal copy-move forgery) in the video. The frames in the video are separated into an overlapping block, and the structural similarity among two consecutive frames are measured to detect the forgery. Zheng et al. [153] presented a technique to detect the frame insertion forgery in the video based on the Block wise Brightness Variance Descriptor (BBVD). They divide the video into a series of frames

[2]The details of test video sequences are available at on Internet via URL: https://sites.google.com/site/multimediaforensic, (STCA, 2013).

and theses frames again partitioned into an overlapping block. BBVD features are extracted and analyzed from each of these blocks to detect the forgery. Wang et al. [128] have presented a technique based on the Consistency of Correlation Coefficients of Gray Values (CCCoGV) to detect frame insertion and deletion forgery in the video. The differences in CCoGV values among adjacent frames of videos are computed to identify the forgery and SVM algorithm is used to distinguish the forged and original video. The authors have also reported the classification accuracy, for a single type of forgery with 25 frame insertions, 100 frame insertions, 25 frame deletions, and 100 frame deletions are 99.22 %, 99.34 %, 94.19 %, and 97.27 % respectively. Whereas classification accuracy for two types of forgery with 25 frame insertions & 25 frame deletions is 96.21 % and with 100 frame insertions & 100 frame deletions is 95.83 %. Yin et al. [145] proposed method using Nonnegative Tensor Factorization (NTF) for the detection and localization of frame insertion/deletion forgery in the video. The method is based on the finding consistency of the time-dimension factor to detect inter-frame forgery. The video is factorized with the use of NTF algorithm, and then the time-dimension factor is extracted from it. At last, The correlation among the extracted elements of the coefficient is compared to detect the forgery. Chittapur et al. [22], have designed a method to detect the region level forgery based on the statistical property of mean and pixel comparison. The temporal difference among each of the video frames is examined to identify and locate the forged region. Tralic et al. [120] presented frame duplication forgery detection method for the video based on Local Binary Patterns (LBP) and Cellular Automata (CA). The video frames are divided into overlapping blocks. Then, the histogram rule is created and applied a CA to every block to detect the forgery. Based on the inconsistency of Quotients of Consecutive Correlation Coefficients of LBPs (QCCoLBPs), Zhang et al. [151] presented video forgery detection algorithm to expose the inter-frame forgery (i.e., frame insertion or deletion). The QCCoLBP is calculated between the neighbouring frames in the video. Then, the Tchebyshev inequality concept is used to detect suspicious abnormal points. The Precision and Recall parameters are taken into consideration to measure the performance of the algorithm. The (Precision, Recall) values for single type of forgery with 25 frame insertions, 100 frame insertions, 25 frame deletions and 100 frame deletions are (98.62 %, 95.33 %), (98.78 %, 94.49 %), (89.27 %, 87.48 %) and (94.31 %, 91.47 %) respectively. Whereas precision, recall values for two types of forgery such as insertion and deletion are 88.16 % and 85.80 % respectively. Singh et al. [105] have suggested a method to identify and locate the frame duplication forgery in the video with the help of block-based features. They divided each frame of video into four sub-blocks (B1, B2, B3, B4) and approximately, nine features from each frame in the form of the mean of a block, ratio and residue for each sub-block are extracted. Then, a lexicographical sort is performed on to the extracted feature to group the similar frames of video. After that Root Mean Square Error (RMSE) value between adjacent frame is calculated, if it is less than a threshold value, then the frames are rejected, and a remaining frame is kept as doubtful. Finally, The correlation between doubtful frames is computed to identify the frame duplication. Pandey et al. [83]-b suggested forgery detection method to expose copy-move forgery in the video frame based on Scale-Invariant Features Transform (SIFT) and K-NN matching algorithms. A compressive sensing technique is proposed by Su et al. [110], to identify moving foreground removal from the video with a static background. They collected the feature difference among the adjacent frames with the use of the Singular Value Decomposition (SVD) algorithm. After that, random projection concept is applied to investigate the features in lower-dimensional space. These features are then clustered using a k-means technique to detect the manipulations. Bagiwa et al. [10] have proposed an approach to detect the chroma key forgery present in the video depends on the correlation among extracted blurring artifact. Chroma key is a kind of splicing forgery in which two videos are combined, with one video's background color becoming transparent to expose another video. They computed cross-correlation between video foreground blocks and background to detect the forgery. Xu et al. [139] have suggested a technique to detect the frame deletion, insertion, and duplication forgery in a video based on the histogram intersection. The correlation coefficients are calculated using the histogram intersection, and then the outliers from it are analyzed to confirm the forgery. Li et al. [65] proposed the method using the uniformity of Quotient of Mean Structural Similarity (QoMSSIM) to detect the frame deletion and insertion forgery in the video. They examined the facts that QoMSSIM are consistent for the original video and it disturbed in case of forged video. QoMSSIM between each of the two frames is calculated and observed for the presence of forgery. They used the SVM to distinguished the original and forged video. The suggested method shows the robustness against recompression and white Gaussian noise. The authors have reported the classification accuracy, for a single type of forgery with 25 frame insertions, 100 frame insertions, 25 frame deletions, and 100 frame deletions are 98.62 %, 98.96 %, 90.72 %, and 94.94 % respectively. Whereas for two types of forgery with 25 frame insertions & 25 frame deletions are 92.27 % and with 100 frame insertions & 100 frame deletions is 92.75 %. Mathai et al. [74] presented the algorithm to detect and localize the content duplication forgery (also called as a temporal copy-move forgery) in video using moment features and cross-correlation concept. The features from the prediction-error array are estimated for every frame-block, and then the normalized cross-correlation is checked to find out the duplication. Yang et al. [140] have proposed approach to detect and localize the frame duplication forgery in a video with the use of a similarity analysis method. The method worked in two steps. In the first step, the features of each frame are collected by using the SVD algorithm. Then, Euclidean distance is computed among the features of every frame with a reference frame. In the second step, the duplications present in the video are identified using random block matching. Liu et al. [71] have proposed the technique to identify the inter-frame forgeries in the video with the use of Zernike Opponent Chromaticity Moments and Coarseness Analysis (ZOCM). The same authors presented a Three-Stage Foreground Analysis And Tracking Algorithm (3FAT) in [72] to identify the blue screen composition video forgery. They exploited irregularities of the contrast and luminance between background and foreground to detect the forgery. In the first step, foreground blocks in a video are extracted using the multi-pass foreground locating method such as GMM. After that, to detect the forged block, A mixture of local features, such as luminance, contrast, *etc*. are used to verify the resemblance of the foreground block and the background. Finally, the forged block in a subsequent frame is monitored with the assistance of a compressive monitoring concept using a quick target search algorithm. Bozkur et al. [15] have introduced the technique to detect the frame duplication forgery and localization of it in video based on forgery line. They divided each frame of video into the non-overlapping sub-blocks, and DCT is applied to each of these sub-blocks. After that, a row vector that contains the averaged DCT values is created from each frame. These row vectors are then binarised to compute a correlation matrix and creates a correlation frame. Finally, hough transform is used on the correlation frame to find forgery line to detect the forgery. Based on the binary features, the technique to detect and locate the frame duplication and frame mirroring forgery in the video is proposed by Ulutas et al. [122]. Firstly, the video is split into the frames, and each frame then transformed into a binary form. The binary features from these frames are extracted to determine the similarity among feature. After that, the Euclidean distance measure is computed for analyzing the similarity among adjacent frames. Then Peak Signal to Noise Ratio (PSNR) values among similar frames are measured to avoid the false duplication. At last, the post-processing operation is applied to enhance performance. The same authors have designed a method to handle the frame duplication forgery present in the video using Bag of Words (BoW) model in [123]. The BoW model is invented here to generate visual words and construct the dictionary from SIFT key points of frames in the video to detect the duplicated parts. A patch-based algorithm to identify and localize the copy-move forgery in the video with the help of Zernike moments features is mentioned in D'Amiano et al. [24]. The similarity analysis-based scheme is developed by Zhao et al. [152], to detect and localize the forgeries like frame deletion, insertion, and duplication with the help of histogram and Speeded Up Robust Features (SURF). In the first module, the HSV (Hue-Saturation-Value) color histogram comparison algorithm is used to detect the forgeries. The SURF and FLANN (Fast Library for Approximate Nearest Neighbors) algorithms are used in the second module to localizes the forgery. Su et al. [111] have suggested a forgery identification method using Exponential-Fourier Moments (EFMs) features to identify the region duplication forgery (also called as copy-move manipulation) in videos. EFMs features are extracted from every block of the current frame and check whether there is a matching pair or not. Then, the Post-Verification Scheme (PVS) is used to eliminate manipulated pairs and locate the forged area in the video frame. At last, an Adaptive Parameter based Fast Compression Tracking (AFCT) method is used for checking the forged areas in the corresponding frames. The proposed method worked efficiently for the forged region with mirroring attack (mirror invariant). Furthermore, the same authors have presented a technique in [109] for detecting the duplication forgery (Copy-Move forgery) in the digital video using Mirror-invariant and Inversion-invariant SIFT (MI-SIFT). The MISFIT algorithm is used to extracts features from the current video frame. Then, the manipulated regions in the current video frame are detected. At last, Spatio-temporal context learning algorithm is created to finds the manipulated regions in the other frames. Moreover, authors have developed another algorithm in [112] to detect the forgery in videos with variable bit-rate compression for the detection of foreground removal (also called as object removal) forgery in the video. They created the Energy Factor to detect forged frames and locate the manipulated region in those frames by developing an adaptive parameter-based visual background extractor (AVIBE). The proposed algorithm is robust against post-processing operation like noise addition, brightness change, shaking screen and water ripples. Wei et al. [136] developed the detection technique based on a multi-scale standardized mutual information to detect inter-frame forgeries such as frame duplication, insertion, and deletion forgery in the video. The crucial features are extracted from the frames, and then the similarity between the adjacent frames is calculated using the relevant measurement function. Based on the correlation coefficients and coefficients of variation, Singh et al. [100] developed two separate algorithms to detect the forgery in videos. The first algorithm extract mean features from each frame and estimate the correlation among the frames to detect the frame duplication forgery. In contrast, the second algorithm estimates the similarity among region within the frames to locate the copy-move forgery. The algorithms are tested on both static and moving background videos. To detect and localize the frame insertion, duplication, and deletion forgery in video Bakaset al. [12] proposed the approach by analyzing the Haralick correlation inconsistency among the frames. The benefit of the proposed approach is that it is independent of GOP size/structure, and the number of frame deletion. Also, it is suitable for both slow-motion static and moving background videos encoded with MPEG-4, XViD, H.264 and H.265

codecs. The authors tested the proposed approach on static as well as a dynamic background video and reported some parametric values such as precision, recall and F1score. In case of video with static background parametric values for frame insertion/deletion detection and localization are PR=85 %, RR=89 %, F1-Score=87 % and PR=95.8 %, RR=94.2 %, F1-Score=94.8 % respectively whereas for frame duplication detection and localization the values are PR=93 %, RR=100 %, F1-Score=96 % and PR=98.8, RR=100 %, F1-Score=99 % respectively. In case of Dynamic background video, parametric values for frame insertion/deletion/duplication detection and localization are PR=95.6 %, RR=82.4 %, F1-Score=88.4 % and PR=99.4 %, RR=97.6 %, F1-Score=98.4 % respectively. Bai et al. [11] presented a technique to identify and locate the TCP and ETS inpainting forgery in video using Spatio-temporal LBP analysis. The proposed method is tested on both static as well as moving background video. However, performance is affected by fast-moving background videos. Aparicio et al. [8] presented a technique to detect and locate the copy-move and frame duplication forgery in video using a block correlation matrix. The block correlation matrix is used to stores both the spatial and temporal information of all the pixels to detect the forgery. Based on texture inconsistency, Saddique et al. [94] proposed a new method to detect the region manipulation forgery in the video. Firstly, the Difference of Consecutive Frames (DOCFs) from the video sequence is calculated. Discriminating features are then extracted via a CCD-DRLBP (Chrominance value of Consecutive Frame Difference and Discriminative Robust Local Binary Pattern) descriptors which is then helpful for the detection and localization of forgery. These extracted features are then supplied to the SVM to identify video clips as authentic or forged one. The proposed approach is robust against the geometric transformation and post-processing operations. However, it is not suitable for the video captured through moving camera. Aloraini et al. [5] have proposed an approach for detecting the object-based forgery (specifically moving object) in the video. The proposed approach divided into three stages such as spatio-temporal filter, sequential analysis and object movement estimation. In spatio-temporal filter stage, the video is divided into frames, and spatial decomposition is applied with the help of Laplacian pyramid.[3] Then the temporal high pass filter is used to detect the edges. The Sequential analysis is the second stage which is used to identify the pixels change in video frames. At last, the forged object estimation is done by summarizing all the pixels change in video frames. Furthermore, The same authors have modified the existing approach based on Sequential and Patch analysis and developed a new approach in [6] for the identification and localization of object removal forgery in the video. In Sequential analysis, video sequences are modelled as stochastic processes and alterations in the parameter during sequence modelling are explored for the detection of forgery. Whereas in Patch analysis, video sequences are modelled as a combination of normal and abnormal patches to identify the distribution of each patch. Finally, the forged regions are localized by observing the movement of the removed objects using abnormal patches. Kharat et al. [51] proposed a two-stage algorithm to identify the frame duplication forgery in MPEG 4 video. The motion vectors for all the frames are determined to classify suspicious frames in the first stage. In the next stage, SIFT features of every frame are calculated to take the final decision to identify duplication forgery. The suggested method works fine for both on compressed and uncompressed videos with different compression rate.

[3] The Laplacian pyramid is a flexible data structure with several appealing features for image/video analysis.

3.2.5 Machine learning-based techniques

The use of machine learning techniques in the area of computer vision encourages the researchers to apply machine learning (ML) and deep learning (DL) models for video forgery detection. These techniques are data-driven (i.e., which need a huge amount of data), and they are capable of automatically learning necessary complex features/artifacts required to detect the forgery in the video. The different types of ML/DL models such as SVM, K-Nearest Neighbour (KNN), Logistic Regression (LR), Linear Discriminant Analysis (LDA), Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Auto Encoder *etc.* are used by the researchers for the detection of forgery in the video. The analysis of video forgery detection techniques based on machine learning is shown in Table 5. Shanableh et al. [95] have presented the machine learning-based approach for the detection of frame deletion forgery in the digital video. They extracted different features such as prediction residuals, quantization scales, percentage of intra-coded macroblocks, and PSNR values from the video. They used machine learning methods such as K Nearest Neighbor (KNN), Logistic Regression (LR) and SVM to detect the deletion forgery from a video. They used 36 MPEG-2 coded videos with Constant Bit Rate (CBR) and Variable Bit Rate (VBR). The presented method works on CBR and VBR encoded video with both fixed GOP and variable GOP length structure. Yao et al. [143] have designed deep CNN model to handle the object-based forgery in a video. They transformed the input video into image patches by using an absolute difference algorithm. Then the training data set is generated, which is labelled as a positive and negative sample of image patches. After that, the five-layer CNN model is trained using the generated training data. They used Caffe deep learning framework [46] to implement the CNN model. The designed model is tested on videos (encoded with H.264/MPEG-4 codec) taken from SYSU-OBJFORG dataset [20]. Long et al. [73] have proposed Convolutional 3D Neural Network (C3D) model to detect and localize the frame deletion or dropping forgery in a video by exploiting the Spatio-temporal relationship in the digital videos. The proposed model is tested and validated on videos are taken from the Yahoo Flickr Creative Commons 100 Million (YFCC100m) [144] and Nimble Challenge 2017 dataset [79]. The proposed model is suitable for the video with stationary and moving background videos. The work by D'Avino et al. [27] used the deep learning model based on autoencoder and RNN to detect the splicing forgery in a video. They extracted frame residual-based features to train the network. The experiment is implemented in TensorFlow using the Adam learning algorithm and tested on a personal dataset, which is available at [34]. The limitation of their model is that it takes too much time to train the deep learning network. Based on the Spatio-temporal consistency, Kono et al. [55] have proposed Convolutional Long Short-Term Memory (ConvLSTM) models to detect the object removal forgery in the video. They used CNN to consider the spatial aspects of the video, whereas RNN is used to consider the temporal aspect of the videos. The method works for both static and dynamic background videos. Hong et al. [39] presented a scheme to delete the frame deletion forgery in HEVC encoded video. They concentrated on the sort of frame changes that occur when the frame is deleted, which create subtle differences between both the coding patterns in the source and the manipulated video. The proposed scheme consists of two parts. In the first part, the useful features from compressed coding information are extracted. The second part uses the classifiers such as LDA, KNN and MLP to check the genuineness of the video. The benefit of this scheme is that it is designed for the video encoded with the latest codec, HEVC. Johnston et al. [48] proposed a framework for localization of region tampering in video

Table 5 Analysis of video forgery detection techniques based on machine learning

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Shanableh et al. [95]	Prediction residuals, Quantization scales, PSNR values and KNN, LR & SVM	Frame Deletion	36 MPEG-2 coded video	Avg. TPR=95 % Avg. FNR=4 %	Exact localization is not done. Handle a single type of forgery.
Yao et al. [143]	CNN	Copy-move or Object Removal	SYSU-OBJFORG [20] (H.264/MPEG-4 encoded video)	PFACC=98.08 % FFACC=88.75 % FACC=96.68 % PR=96.5 % RR=90 % F1-Score=93.26 %	Localization is missing. Not suitable for moving background videos with high resolution & high bitrate. Computation cost is high.
Long et al. [73]	CNN	Frame Deletion	2650 forged videos created (Use Yahoo Flickr Creative Commons 100 Million (YFCC100m) [144] and Nimble Challenge 2017 [79].	DA=99.83 %	Need high computational time. Work with fixed-length GOP video. Work on single-shot video & for a single type of forgery.
D'Avino et al. [27]	Autoencoder and RNN	Splicing	Created own dataset of 10 videos	ROC	Need more computational time. Work with a static background & fixed-length GOP video.
Kono et al. [55]	Spatio-temporal consistency, ConvLSTM	Object Removal	CDnet 2014 [135]	AUC= 0.977, Equal-Error-Rate=0.061	GOP size detail is not mentioned. Design for a single type of forgery.

Table 5 (continued)

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Hong et al. [39]	Compressed coding information, LDA using KNN & MLP	Frame Deletion	Personal Dataset	For LDA + KNN: PR=79.5 %, RR= 89.5 %, DA=82.3 %, F1-Score=84.3 % & For MLP:PR=86.5 %, RR=91.7 %,DA=88.3 %, F1-Score=88.8 %	Work with fixed GOP size & static video.
Johnston et al. [48]	Compression parameters & CNN	Region Manipulation, Copy-Move & Splicing	Videos from [4, 27, 90]	F1-Score	Not able to detect multiple manipulations. Addition features need to be examined. Work with fixed GOP size.
Zampoglou et al. [147]	DCT, quantization error, Deep CNN	Inter-frame forgeries and Region Manipulation	Created own dataset (Videos taken from [79])	Nil	Localization is missing. Proper annotation is needed for forged video. Localization is missing. Proper annotation is needed for forged video. Work with a static background & fixed-length GOP video.

using features learned from original video contents. They used CNN to estimate the compression parameters like quantization scale, deblock filter setting and intra/inter-frame type. Zampoglou et al. [147] presented a technique to detect the double quantization, frame insertion and region manipulation in video using Deep CNN. They designed two forensics filter one is based on DCT and other is based on quantization error. The filter outputs are then given to the Deep CNN to differentiate original and forged video.

4 Anti-forensics techniques

Video anti-forensic techniques have been developed to deceive forensic investigation by removing or concealing traces left after the forgery. Although forensic techniques are useful in identifying digital manipulations in videos, most of them could fail if a forger uses anti-forensic approach. The anti-forensic techniques work on the principle that if someone removes or reduces the traces left over after the manipulations in the video, which itself leads to other evidence and that need to be further investigated to identify the forgery. Stamm et al. [108] concentrated on the periodical re-compressed artifacts left after the frame insertion and deletion forgery. They designed the anti-forensic technique by identifying the P-frame prediction error in a manipulated digital video. Furthermore, the counter anti-forensic technique has also been designed to make a comparative analysis between the actual predicted error and a predicted error acquired from the video. Su et al. [113] presented an anti-forensic method where the inter-frame relationships of coding modes in adjacent frames are analyzed to determine whether the intra-prediction can be applied during the re-encoding process of the tampered video. After re-encoding, the coding parameters and the bit-rates are also examined to predict the targeted distribution of quantization indices to detect the tampered video. Kang et al. [50] modified the frame deletion detection methodology in [70] and proposed new methodology which can also detect the frame insertion forgery in the video. The authors also designed an anti-forensic method based on the analysis of P frame prediction error for the detection frame deletion forgery. Besides, the counter anti-forensic approach for frame deletion forgery have also been proposed, where the predicted error is estimated and then after it is compared with the stored prediction error. Yao et al. [142] focused on the inter-frame interpolation as an anti-forensics operation to identify frame deletion forgery in the video. The method is tested on video encoded with H.264 and H.265 codec with GOP default size of 250. The analysis of anti-forensics techniques for video forgery detection is shown in Table 6.

5 Deepfake detection

Deepfakes are media that use the machine learning to take a person in an actual photo or video and replace them with someone else's identity. Deepfakes were used in porn pictures and videos to swap faces of politicians or celebrities. Hence, deepfake video can be misused to trigger political or religious instability to fool the public and affect election campaign results or disrupt financial markets by creating fake news stories [78]. Figure 18 shows the example for deepfake video wherein the original face is replaced from a new one. The analysis of deepfake detection techniques is shown in Table 7.

Li et al. [64] examined the fact that a normal human would usually blink somewhere among 2-10 seconds, and it would take 0.1-0.4 seconds for every blink. Authors also noted that blinking rates in deepfake video are relatively lower than those in normal videos. Based

Table 6 Analysis of anti-forensics techniques

Ref.	Features/ Methods	Forgeries Identified	Datasets	Performance Parameters	Limitations
Stamm et al. [108]	*P*-frame prediction error	Frame Deletion & Insertion	36 QCIF format Videos from [126, 138]	DA=85 % & FPR=15 %	Not useful for finding the location of the deleted frame. Work well with fixed GOP size.
Su et al. [113]	Recorded MB-types	Frame Deletion	CIF videos encoded with H.264 codec	Nil	Work with fixed GOP size. Localization is not done.
Kang et al. [50]	*P*-frame prediction error	Frame Deletion and Insertion	32 QCIF videos in YUV-uncompressed format from [126, 138]	DA=100 % FPR=6.3 %	Use a fixed GOP structure. Not appropriate for moving background video. Localization is not done.
Yao et al. [142]	Inter-frame interpolation	Frame Deletion	100 CIF video sequences encoded with H.264 codec	DA=90 %	Work with Static background video. Useful for frame deletion forgery only.

Fig. 18 Deepfake example [25]

on these physiological signal (such as eye blinking), they proposed a Long-term Recurrent Convolutional Neural Networks (LRCN) model to detect the deepfake video. The set of eye sequences are provided as an input to the LRCN model, which consist of three stages such as 1) feature extraction 2) sequence learning 3) state prediction. The same authors proposed deep learning-based model in [63] to detect the deepfake videos with the help of face wrapping artifact. The CNN models such as such as VGG16 [99], ResNet152, ResNet101 and ResNet50 [38] are used to detect the deepfake forgery. The PRNU analysis is adopted by Koopman et al. [56] to expose the deepfake detection in a video. They divide the video into frames and faces are cropped out from those frames. The extracted faces are then divided into groups and PRNU calculated for each of these groups. After that, the mean normalized cross-correlation score is calculated to distinguish deepfakes from authentic videos. Guera et al. [37] explored the intra-frame frame and inter-frame consistency between video frames and developed the temporal-aware pipeline approach using CNN and LSTM model. The frame-level features are extracted using CNN, which are then fed to the LSTM model to detect the deepfake video. The proposed model is tested on 300 deepfake videos with an average accuracy of 96.96 %. Afchar et al. [1] proposed a MesoNet deep learning network to observe the mesoscopic properties of images/frames for detecting the forged video of faces. They evaluate the proposed deep network on fake video dataset with an average detection rate of 98 %. To identify the deepfake video, a Recurrent Convolutional Network (RCN) model is suggested by Sabir et al. [92]. The model is based on the integration of the CNN features with DenseNet [44] and the gated recurrent unit cells [23] to analyze the temporal correlation across frames. The suggested model is tested on the FaceForensics++ dataset [91], that consist of 1,000 videos. Yang et al. [141] presented a deepfake detection method by analyzing the differences between 3D head poses containing head orientation and position. The extracted artifacts are given to the SVM classifier to get the detection result. Nguyen et al. [77] suggested capsule networks that identify the manipulation in images and videos. They used VGG-19 network [99] to extract the latent features from video frame and then fed it to the capsule networks (which is based on dynamic routing algorithm [93]) for classification. Zhang et al. [149] have presented a novel transfer learning-based technique to identify the deepfake forgery in the video. They used two neural network model such as

Inception-v3[4] and MobileNet V1 [40] to detect the deepfake video. Amerini et al. [7] presented a technique to expose the deepfake detection in video using optical flow coefficients and CNN classifier. Firstly, they divide the video into frames, and then optical flow coefficients among all these frames are extracted. Finally, the extracted features fed to the CNN model to identify the original or fake video.

6 Video forgery datasets

In this section, the analysis of existing available video forgery dataset is studied and analyzed. Table 8 shows the analysis of video forgery datasets. Qadir et al. [86] have created another video dataset for testing video forgery detection technique named as Surrey University Library for Forensic Analysis (SULFA). It consists only copy-move type of forgery-based videos. The SULFA dataset consists of 150 videos collected from static cameras, and it is available online at [117]. Each video in a dataset is 10 seconds long, with a frame rate of 30 fps and has a resolution of 320 × 240. SYSU-OBJFORG is one of the forged video datasets, which comprises of 100 original video footages and 100 forged video footages developed by Chen et al. [20]. These video sequences are of 11 seconds long, with a resolution of 1280 × 720, compressed by H.264/MPEG-4 codec with a bit rate of 3 Mbit/s and has a frame rate of 25 fps. REWIND forged video dataset is created by Bestagini et al. [13]. They used SULFA dataset [86] to create their dataset. This dataset consists of 10 original and 10 forged videos which are having a resolution of 320 × 240 pixels with a framerate of 30 fps and compressed with MJPEG and H264codec. REWIND dataset contains the differences between the frames of the original sequences and the forged sequences, which is useful in video forgery detection. The dataset is available at [88]. Ulutas et al. [123] have created a dataset which consists of 31 forged videos (with both static and moving background videos) with frame duplication forgery. They perform the manipulation on 25 videos are taken from SULFA dataset [86] and 6 videos from different movie scenes using virtual dub software. The dataset is available online at [26]. D'Amiano et al. [24] have created a dataset which consists of 15 forged videos with copy-moves forgery (forged videos with 10 additives and 5 occlusives). They used After Effects Pro tool to perform the forgery in the video. The dataset is available online at [35]. Davino et al. [27] have created the dataset which contains the forged video with splicing forgery. This dataset contains 10 forged videos along with the 10-original video. The Adobe After Effects CC tool is used to perform the forgery in the video. The dataset is available at [34]. Al-Sanjary et al. [4] created a Video Tampering Dataset (VTD) which contain manipulated videos which are used for testing the performance of video forgery detection technique. Videos are collected from YouTube and networking websites. The VTD includes 33 videos, categorized among three types of forgeries such as Splicing forgery, Copy-Move forgery, and Swapping-Frames. The length of each video is of 16 seconds, with a resolution of 1280 × 720, and a rate of 30 frames per second. Their dataset is available at [125]. Ardizzone et al. [9] have created datasets of tampered videos by cloning the objects (copy-move forgery) from a video sequence. Also, they applied various transformations on tampered videos such as Scaling, Shearing, Rotations, Flipping, Luminance and RGB. They gathered different videos from SULFA [86] and CANTATA [16] video datasets for the scenario related to traffic control

[4] Inception V3 [118] is a CNN that is trained on over a million of images from ImageNet dataset.

Table 7 Analysis of deepfake detection techniques

Ref.	Feature/ Methods	Datasets	Details	Limitations
Li et al. [64]	Eye blinking	49 real and 49 deepfake videos	LRCN is used to learn the temporal patterns of eye blinking.	Require a huge amount of images with closed eye. Improvement is required on the dynamic pattern of blinking.
Li et al. [63]	Face warping artifacts	UADFV dataset consist of 49 real and 49 fake videos [64] & Deepfake TIMIT [57]	VGG16 [99], ResNet152, ResNet101 and ResNet50 [38] are used.	Not robust against multiple video compression.
Koopman et al. [56]	PRNU Analysis	10 real16 deepfake video	Explore the PRNU patterns between the real and deepfake videos.	Testing is done on a small dataset, so there is a need to test the work on massive datasets.
Guera et al. [37]	Intra-frame and temporal inconsistencies	600 videos obtained from websites	CNN and LSTM are used.	Need to improve the robustness of technique against the unseen manipulations.
Afchar et al. [1]	MesoNet	Created own dataset of 175 forged videos	Meso-4 and MesoInception-4 are used to investigate deepfake videos at the mesoscopic analysis level.	Due to the use of macroscopic features, the model becomes complicated to understand.
Sabir et al. [92]	Using spatio-temporal features	FaceForensics++ data set [91]	Temporal inconsistencies across adjacent frames are analyzed using RCN.	Leads to overfitting problem due to limited sample in FaceForensics++ dataset.
Yang et al. [141]	Head Poses	UADFV [64] consist of 49 real and 49 fake videos and MFC datasets [36]	Features are obtained using 68 landmarks of the face area. SVM classifier is used.	Provide average accuracy.
Nguyen et al. [77]	Capsule-forensics	Deepfake dataset [1]	VGG-19 [99] network is used for feature extraction, and Capsule networks are used for classification. For face-swapping detection DA =95.93 % (at frame level) DA = 99.23 % (at video level).	Not robust against an intentional attack such as Noise addition.

Table 7 (continued)

Ref.	Feature/ Methods	Datasets	Details	Limitations
Zhang et al. [149]	Inception-v3 and MobileNet V1	FaceForensics dataset [90]	DA=94.9 %	Robust testing is needed to improve the reliability of the results.
Amerini et al. [7]	Optical Flow Coefficients and CNN classifier	FaceForensics++ dataset [91]	For VGG16: DA= 81.61 % For ResNet50: DA= 75.46 %	Testing against more deepfake datasets is needed to check the reliability of the optical flow field.

Table 8 Analysis of video forgery datasets [FPS: Frame per Second]

Ref.	Dataset Name	Forgery Present	Video Source	Video count	Format/ Codec	FPS	Resolution	Camera Type
Qadir et al. [86]	SULFA	Copy-Move	Canon SX220, Nikon S3000, Fujifilm S2800HD	150	MOV & AVI (codec H.264, MJPEG)	30	320 × 240	Static
Chen et al. [20]	SYSU-OBJFORG	Object based forgery (Adding or removing the moving object)	Commercial Surveillance Cameras	110	H.264/MPEG-4 encoded	25	1280 × 720	Static
Bestagini et al. [13]	REWIND	Copy-Move	SULFA [86]	10	MOV & AVI (codec H.264, MJPEG)	30	320 × 240	Static
Ulutas et al. [123]	Test Database	Frame Duplication	SULFA [86] & Different movie scene	31	MPEG-4	-	Variable	Static & Dynamic
D'Amiano et al. [24]	GRIP dataset	Copy-Move (Additive & occlusive)	Internet	15	AVI	30	Variable	Static
D'Avino et al. [27]	GRIP dataset	Splicing	YouTube & Internet	10	AVI (codec H.264)	30	720 × 1280	Static
Al-Sanjary et al. [4]	VTD	Splicing forgery, Copy-Move forgery, & Swapping-Frames	Internet	33	AVI	30	1280 × 720	Static & Dynamic
Ardizzone et al. [9]	Not mentioned	Copy-Move	SULFA [86] & CANTATA [16]	160	AVI & MP4	25 & 30	960 × 540 640 ×360 320 × 240	Static & Dynamic

and parking surveillance. Their dataset contains 160 forged videos with an average duration of 30 cloned frames.

7 Generalized architecture of passive video forgery detection

Video forgery detection using passive techniques are binary classification techniques. The main aim of these techniques is to classify given videos into two classes, such as original and forged videos. Most of the existing passive forgery detection techniques, first extract distinct features from videos, then select an appropriate classifier and train it using the extracted feature set to classify the videos. Few such techniques are proposed in Chen et al. [20], Aghamaleki et al. [2], Aghamaleki et al. [3], Hsu et al. [41], Ravi et al. [87], Kancherla et al. [49], Wang et al. [129] Tan et al. [119], Chen et al. [19], Lin et al. [69], Wang et al. [128], Li et al. [65], Shanableh et al. [95], Yao et al. [143], Long et al. [73], D'Avino et al. [27], Sabir et al. [92], Yang et al. [141], Guera et al. [37] and Nguyen et al. [77]. The generalized architecture for passive video forgery detection technique is shown in the Fig. 19 which consist of the following important stages:

- Pre-processing: - The main objective of pre-processing is an enhancement of the digital video frames that suppresses from unnecessary alteration or improves some features crucial for later processing. Before the feature extraction stage, some important operations have to be performed on the video, like RGB to gray conversion, DWT or DCT transformation and cropping to optimizes the classification performance.
- Feature Extraction: - This stage starts with a set of calculated data and builds resultant values which are called as features that considered being relevant and non-redundant. A collection of features shall extract for every class of video frame that is used to differentiate it from other classes. In digital video analysis, feature extraction obtains the useful artifact from a video which will be helpful for further investigation.
- Feature Pre-processing: - The use of this module is to decrease the feature dimensionality without significantly reducing the efficiency of classification.
- Forgery Detection Technique: - The main aim of this stage is to apply certain techniques on extracted and pre-processed features for detecting the forgery in the digital video.
- Classification: - The prime use of this module is to analyze to which of the class a new inspection fits in, with the use of a training set of video contents containing observations

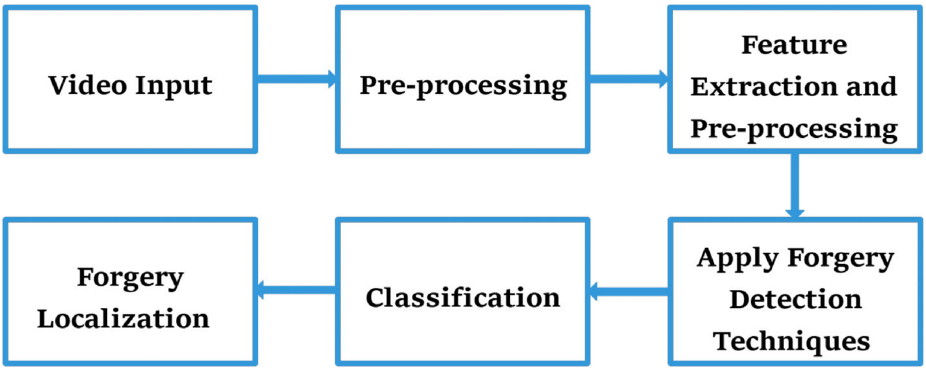

Fig. 19 Generalized architecture for passive video forgery detection

whose class is known. Based on the extracted collection of chosen features, the suitable classifier is designed to make a distinction between the original and the forged video.

- Forgery Localization: - The main target of this stage is to locate the exact place of the forgery present in the video.

8 Discussion and new challenges in video forgery detection

Based on the study of various passive video forgery detection techniques invites several merits and demerits illustrated in Table 9.

Table 9 Merits and demerits of various passive video forgery detection techniques

Techniques	Merits	Demerits
Compression Artifacts	Almost all the video present over the Internet is in a compressed format to solve the storage problem. Specifically designed for compressed videos to detect/localize both inter and intra-frame forgery. Also suitable for the detection of double compressed video.	Performance relies on the video codecs used for compression. Change in video bit-rates and quantization scale ratio affect the performance.
Noise Artifacts	Noise is an essential feature or a clue in the video forensics. Almost every video contains several sorts of noise. So, it will be easy for the researcher to work on this feature directly.	The video noise has clearly changed over the last 15 years; it's been a challenging task for the researcher to create a new methodology as of for the new type of noise.
Motion Features	Motion is an essential feature in video forgery detection. Detect/Localize both inter and intra-frame forgery, but it is mostly suitable for inter-frame forgery. Especially the optical flow algorithm is one of the frequently used algorithms to detect the forgery in the video.	The performance of the detection techniques may affect due to the speed of the video and background of the video.
Statistical Features	Almost every video consists of statistical feature such as pixel correlation, geometric and physical properties, so it is one of the most widely used techniques for forgery detection. Detect/Localize both inter and intra-frame forgery	Need to study several algorithms for a different type of statistical features. Computational overhead is high due to the complex correlation calculations.
Machine Learning-Based Techniques	Emerging techniques in video forgery detection domain. Less human interaction is needed due to the use of ML/DL models. No need to extract the hand-crafted aritifacts/features from the forged video. Also, useful for deepfake detection in video.	Somehow costly as it requires resource like high-end GPU. Need a huge dataset and required more computational power. Currently useful in identifying specific forgery.

It is observed from the study on various existing passive video forgery detection techniques, for a particular scenario, the suitable method for detecting the forgery relies on following essential parameters, such as.

- **Compression:** The performance of most of the video forgery detection techniques discussed in the literature relies on the video codecs such as H.264, MPEG-4, MPEG-3, MPEG-2, and MPEG-1 used for compression. Compression artifacts based techniques may fail in uncompressed forged videos. It is recognized that the forgery detection accuracy of many of the existing techniques decreases with the increase in compression ratio. Also, it is affected by the change in video bit-rates and quantization scale ratio. In most of the cases, compression artifacts present in the video degrades the performance of the detection system. Many of the techniques proposed so far able to detect the forgeries in video compressed with specific codec only. Video recompression using the same encoding parameters and forgery identification in highly compressed videos are some of the issues that need to be addressed.
- **GOP's Structure:** The most usually used video encoder such as H.264/AVC uses adaptive GOP's structure in the current scenario where the GOP size will expand up to 250 frames depending on the video content changes. Many of the mentioned techniques work well for GOP with a fixed structure size, and quite a few them are useful for detecting the forgery in variable GOP structure videos and are unable to detect the deletion of a complete GOP or multiples of GOP's.
- **Noise:** Since the video noise has clearly changed over the last 15 years; it's been a challenging task for the researcher to create a new methodology as of for the new type of noise. Also, it is observed that the noise present in the video affect the performance of the detection system.
- **Video Background:** Many recent forgery detection techniques designed so far are capable of detecting the forgery in a video with a static background (i.e., not suitable for the video with dynamic or moving background). Exceptionally few techniques are developed to expose the forgery in video with a moving background, so it is another issue for researchers to work on it.
- **Detection and Localization of Forgery:** Most of the stated techniques deal with the identification and localization of a single type of forgery in the video. At the same time, they are not capable of examining multiple forgeries present in the video. Splicing, Frame replication, upscale crop, and frame mirroring are a different kind of forgeries in the digital video, which are not much explored.
- **Video Frame Count:** Most of the present techniques are dependent on the numbers of frame inserted, deleted or duplicated in case of detection of inter-frame forgery. Also, these techniques are not able to detect the forgery in the video when the video frame count is less than a certain threshold.
- **Video Quality and Length of the Video:** Many video forgery detection techniques have designed only for low resolution and short length videos. Due to which there is an extended scope for the researchers to develop a better method to detect and localize the forgery in long length videos.
- **Video Forgery Datasets:** The foremost concern of existing techniques discussed in the literature is the lack of video forgery datasets to perform comparative experimental analysis. The current datasets mostly consist of videos with a single type of forgeries such as copy-move, splicing, and frame duplication, also it mostly contains the forged videos with stationary background only. Very few datasets reviewed in the literature consist of a forged video with a moving background. Presently no such video forgery

Table 10 Summarization of video forgery detection techniques (A: Copy-Move, B: Splicing, C: Region Manipulation (Object insertion or deletion), D: Frame Insertion, E: Frame Deletion, F: Frame Duplication, G: Frame Replication, H: TCP & ETS Inpainting, I: Upscale Crop, J: Mirror Invariant, K: Detection, L: Localization, M: Fixed Size GOP, N: Variable size GOP, O: Video with Static Background & P: Video with Moving Background

Ref.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Wang et al. [130]	×	×	×	✓	✓	×	×	×	×	×	✓	×	✓	×	✓	×
Wang et al. [133]	×	×	×	✓	✓	×	×	×	×	×	✓	✓	✓	×	✓	×
Subramanyam et al. [115]	✓	×	×	×	×	✓	×	×	×	×	✓	×	✓	×	✓	×
Subramanyam et al. [116]	×	×	✓	×	×	×	×	×	×	×	✓	×	✓	×	✓	×
Labartino et al. [59]	×	×	✓	×	×	×	×	×	×	×	✓	×	✓	×	✓	×
Gironi et al. [33]	×	×	×	✓	✓	×	×	×	×	×	✓	✓	✓	✓	✓	×
Liu et al. [70]	×	×	×	×	✓	×	×	×	×	×	✓	×	✓	×	✓	×
Aghamaleki et al. [2]	×	×	×	✓	✓	×	×	×	×	×	✓	✓	✓	✓	✓	×
Aghamaleki et al. [3]	×	×	×	✓	✓	×	×	×	×	×	✓	✓	✓	✓	✓	×
Fadl et al. [29]	×	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	×
Mondaini et al. [76]	✓	×	✓	✓	×	×	×	×	×	×	✓	×	✓	×	✓	×
hsu et al. [41]	×	×	×	×	×	×	×	✓	×	×	✓	✓	✓	×	✓	×
Kobayashi et al. [53]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Kobayashi et al. [54]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
chetty et al. [21]	✓	×	×	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Hyun et al. [45]	×	×	×	×	×	×	×	×	✓	×	✓	✓	✓	×	✓	×
Ravi et al. [87]	✓	×	×	×	✓	×	×	×	×	×	✓	×	✓	×	✓	×
Pandey et al. [83]-a	×	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	×
Hu et al. [42]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Singh et al. [102]	×	✓	×	×	×	×	×	×	✓	×	✓	✓	✓	×	✓	×
Singh et al. [101]	✓	×	×	×	×	×	×	×	×	×	✓	✓	✓	✓	✓	✓
Fayyaz et al. [31]	×	×	×	×	×	×	×	✓	×	×	✓	✓	✓	×	✓	×
Wang et al. [131]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Su et al. [114]	×	×	×	×	✓	×	×	×	×	×	✓	✓	✓	×	✓	×
Dong et al. [28]	×	×	×	✓	✓	×	×	×	×	×	✓	✓	✓	✓	✓	×
Kancherla et al. [49]	×	×	×	×	×	×	×	✓	×	×	✓	✓	✓	×	✓	×
Li et al. [61]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Bestagini et al. [13]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Chao et al. [17]	×	×	×	✓	✓	×	×	×	×	×	✓	×	✓	×	✓	×
Wang et al. [134]	×	×	×	✓	✓	✓	×	×	×	×	✓	✓	✓	×	✓	×
Feng et al. [32]	×	×	×	×	✓	×	×	×	×	×	✓	✓	✓	✓	✓	×
Wu et al. [137]	×	×	×	×	✓	✓	×	×	×	×	✓	✓	✓	✓	✓	×
Wang et al. [129]	×	×	×	✓	✓	×	×	×	×	×	✓	×	✓	×	✓	×
Tan et al. [119]	×	×	✓	×	×	×	×	×	×	×	✓	×	✓	×	✓	×
Bidokhti et al. [14]	✓	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	×
Zhang et al. [150]	×	×	×	×	✓	✓	×	×	×	×	✓	✓	✓	×	✓	×
Yu et al. [146]	×	×	×	×	✓	×	×	×	×	×	✓	✓	✓	✓	✓	×
Chen et al. [20]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×

Table 10 (continued)

Ref.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Singh et al. [103]	×	×	×	✓	✓	×	✓	×	×	×	✓	✓	✓	×	✓	×
Kingra et al. [52]	×	×	×	✓	✓	✓	×	×	×	×	✓	✓	✓	×	✓	×
Sitara et al. [107]	×	×	×	✓	✓	✓	✓	×	×	×	✓	✓	✓	✓	✓	×
Pu et al. [85]	×	×	×	×	✓	×	×	×	×	×	✓	✓	✓	×	✓	×
Wang et al. [132]	✓	×	×	×	×	✓	×	×	×	×	✓	×	✓	×	✓	×
Zhang et al. [148]	×	×	×	×	×	×	×	✓	×	×	✓	×	✓	×	✓	×
Chen et al. [19]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Hu et al. [43]	×	×	×	×	×	✓	×	×	×	×	✓	×	✓	×	✓	×
Lin et al. [61]	×	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	×
Liao et al. [66]	×	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	×
Lin et al. [67]	×	×	×	×	×	×	×	✓	×	×	✓	✓	✓	×	✓	×
Lin et al. [68]	×	×	×	×	×	×	×	✓	×	×	✓	✓	✓	×	✓	✓
Li et al. [60]	×	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	×
Zheng et al. [153]	×	×	×	✓	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Wang et al. [128]	×	×	×	✓	✓	×	×	×	×	×	✓	✓	✓	×	✓	×
Yin et al. [145]	×	×	×	✓	✓	×	×	×	×	×	✓	✓	✓	×	✓	×
Chittapur et al. [22]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Tralic et al. [120]	×	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	×
Zhang et al. [151]	×	×	×	✓	✓	×	×	×	×	×	✓	×	✓	×	✓	×
Singh et al. [105]	×	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	✓
Pandey et al. [83]-b	✓	×	×	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Su et al. [110]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Bagiwa et al. [10]	×	✓	×	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Xu et al. [139]	×	×	×	✓	✓	✓	×	×	×	×	✓	×	✓	×	✓	×
Li et al. [65]	×	×	×	✓	✓	×	×	×	×	×	✓	×	✓	×	✓	×
Mathai et al. [74]	✓	×	×	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Yang et al. [140]	×	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	×
Liu et al. [71]	×	×	×	✓	✓	✓	×	×	×	×	✓	×	✓	×	✓	×
Liu et al. [72]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Bozkurt et al. [15]	×	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	×
Ulutas et al. [122]	×	×	×	×	×	✓	×	×	×	✓	✓	✓	✓	×	✓	×
Ulutas et al. [123]	×	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	✓
D’Amiano et al. [24]	✓	×	×	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Zhao et al. [152]	×	×	×	✓	✓	✓	×	×	×	×	✓	✓	✓	×	✓	×
Su et al. [111]	✓	×	×	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Su et al. [109]	✓	×	×	×	×	×	×	×	×	✓	✓	✓	✓	×	✓	×
Su et al. [112]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	✓
Wei et al. [136]	×	×	×	✓	✓	✓	×	×	×	×	✓	✓	✓	×	✓	×
Singh et al. [100]	✓	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	✓
Bakas et al. [12]	×	×	×	✓	✓	✓	×	×	×	×	✓	✓	✓	✓	✓	✓
Bai et al. [11]	×	×	×	×	×	×	×	✓	×	×	✓	✓	✓	×	✓	✓
Aparicio et al. [8]	✓	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	×
Saddique et al. [94]	✓	✓	×	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×

Table 10 (continued)

Ref.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Aloraini et al. [5]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Aloraini et al. [6]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Kharat et al. [51]	×	×	×	×	×	✓	×	×	×	×	✓	✓	✓	×	✓	×
Shanableh et al. [95]	×	×	×	×	✓	×	×	×	×	×	✓	✓	✓	✓	✓	×
Yao et al. [143]	✓	×	✓	×	×	×	×	×	×	×	✓	×	✓	×	✓	×
Long et al. [73]	×	×	×	×	✓	×	×	×	×	×	✓	✓	✓	×	✓	✓
D'Avino et al. [27]	×	✓	×	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Kono et al. [55]	×	×	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	×
Hong et al. [39]	×	×	×	×	✓	×	×	×	×	×	✓	✓	✓	×	✓	×
Johnston et al. [48]	✓	✓	✓	×	×	×	×	×	×	×	✓	✓	✓	×	✓	✓
Zampoglou et al. [147]	×	×	✓	✓	×	×	×	×	×	×	✓	×	✓	×	✓	×
Stamm et al. [108]	×	×	×	✓	✓	×	×	×	×	×	✓	×	✓	×	✓	×
Su et al. [113]	×	×	×	×	✓	×	×	×	×	×	✓	×	✓	×	✓	×
Kang et al. [50]	×	×	×	✓	✓	×	×	×	×	×	✓	×	✓	×	✓	×
Yao et al. [142]	×	×	×	×	✓	×	×	×	×	×	✓	✓	✓	×	✓	×

dataset is publicly available on the Internet which includes of inter-frame forgeries such as frame insertion, frame deletion, and frame shuffling. Hence, there is ample scope for the researchers to create the forged video dataset for other types of forgery with a moving background.

– **Computational Time:** It is the primary task for researchers to reduce the high computational time needed to detect and locate forgery in the video.
– **Post-processing operations:** The most of forgery detection techniques presented in the survey has not addressed the robustness against post-processing operations such as intentional noise addition, compression and brightness change.
– **Use of Machine Learning/Deep Learning:** Very few techniques are developed so far, which make the use of machine learning methods, especially deep learning. The immense scope is there for the researchers to work with different types of ML/DL models for the detection of both inter/ intra frame forgery in the video. The use of ML/DL models in the area of video forgery detection encourages the researchers to design the automated technique for forgery detection.
– **Inadequate Anti-forensic and Deepfake Detection Strategies:** Very few anti-forensics techniques are developed so far to expose the forgery in the video. Especially most of the techniques designed can handle frame deletion forgery only. So, it has become a great chance for the researchers to explore the anti-forensic strategies for other types of forgery. Furthermore, deepfake detection in the video is one of the hot areas for further research in video forensics domain.
– **Audio aspect in Video:** Although the visual contents of video help us in legal matters at the same time, it is impossible to ignore the role of audio in making the decision. All the existing forgery detection technique proposed so far only focused on visual content, i.e., no attention has been given to the audio component of digital video.

We believe that this study will enable researchers working in the field of video forgery detection to find new useful approaches and ideas. The detail summarization of video forgery detection techniques is presented in Table 10.

9 Conclusions

This paper presented a comprehensive analysis of passive video forgery detection techniques. The detailed analysis of passive video forgery detection techniques is performed in terms of features/method used, forgery identified, datasets used, performance parameters along with their limitations. The emerging topic, such as anti-forensics strategies and deepfake detection in the video have also been discussed. Furthermore, the standard benchmark datasets related to video forgery have been reviewed. Some of the critical challenges which can contribute to significant research in this field has also been mentioned. Although the researchers have proposed several techniques for passive video forgery detection, still there is a necessity to introduce some new techniques which can overcome the points discussed in Section 8. It is observed that most of the existing video forgery detection techniques deal with identifying a single type of forgery and are unable to deal with multiple forgeries. Also, most of the current techniques are dependent on the size of GOP's structure, codec used for video compression, compression rate, noise, size/length of the video, video frame count and background of the video. Very few techniques are designed so far that can detect the forgeries in the video with the help of machine/deep learning. Anti-forensic and deepfake detection in the video is the new aspects that need to be explored more. This survey will be helpful for the research fraternity to improve passive video forgery detection techniques with new ideas.

References

1. Afchar D, Nozick V, Yamagishi J, Echizen I (2018) Mesonet: a compact facial video forgery detection network. In: 2018 IEEE international workshop on information forensics and security (WIFS). IEEE, pp 1–7
2. Aghamaleki JA, Behrad A (2016) Inter-frame video forgery detection and localization using intrinsic effects of double compression on quantization errors of video coding. Signal Process Image Commun 47:289–302
3. Aghamaleki JA, Behrad A (2017) Malicious inter-frame video tampering detection in mpeg videos using time and spatial domain analysis of quantization effects. Multimedia Tools Appl 76(20):20691–20717
4. Al-Sanjary OI, Ahmed AA, Sulong G (2016) Development of a video tampering dataset for forensic investigation. Forensic Sci Int 266:565–572
5. Aloraini M, Sharifzadeh M, Agarwal C, Schonfeld D (2019) Statistical sequential analysis for object-based video forgery detection. Elect Imag 2019(5):543–1
6. Aloraini M, Sharifzadeh M, Schonfeld D (2020) Sequential and patch analyses for object removal video forgery detection and localization. IEEE Trans Circ Syst Vid Technol
7. Amerini I, Galteri L, Caldelli R, Del Bimbo A (2019) Deepfake video detection through optical flow based cnn. In: Proceedings of the IEEE international conference on computer vision workshops, pp 0–0
8. Aparicio-Díaz E, Cumplido R, Gort P, Lázaro M, Feregrino-Uribe C (2019) Temporal copy-move forgery detection and localization using block correlation matrix. J Intel Fuzz Sys 36(5):5023–5035
9. Ardizzone E, Mazzola G (2015) A tool to support the creation of datasets of tampered videos. In: International conference on image analysis and processing. Springer, pp 665–675
10. Bagiwa MA, Wahab AWA, Idris MYI, Khan S, Choo KKR (2016) Chroma key background detection for digital video using statistical correlation of blurring artifact. Digit Invest 19:29–43
11. Bai S, Yao H, Ni R, Zhao Y (2019) Detection and localization of video object removal by spatio-temporal lbp coherence analysis. In: International conference on image and graphics. Springer, pp 244–254
12. Bakas J, Naskar R, Dixit R (2019) Detection and localization of inter-frame video forgeries based on inconsistency in correlation distribution between haralick coded frames. Multimedia Tool Appl 78(4):4905–4935

13. Bestagini P, Milani S, Tagliasacchi M, Tubaro S (2013) Local tampering detection in video sequences. In: 2013 IEEE 15Th international workshop on multimedia signal processing (MMSP). IEEE, pp 488-493
14. Bidokhti A, Ghaemmaghami S (2015) Detection of regional copy/move forgery in mpeg videos using optical flow. In: 2015 The international symposium on artificial intelligence and signal processing (AISP). IEEE, pp 13–17
15. BOZKURT I, Bozkurt MH, Ulutaş G (2017) A new video forgery detection approach based on forgery line. Turkish J Elect Eng Comput Sci 25(6):4558–4574
16. CANTATA (Accessed 2 Nov 2019) Dataset [Online]: http://www.multitel.be/cantata/
17. Chao J, Jiang X, Sun T (2012) A novel video inter-frame forgery model detection scheme based on optical flow consistency. In: International workshop on digital watermarking. Springer, pp 267–281
18. Chen CC, Chen LY, Lin YJ (2017) Block sampled matching with region growing for detecting copy-move forgery duplicated regions. J Inf Hiding Multimed Signal Process 8(1):86–96
19. Chen R, Dong Q, Ren H, Fu J (2012) Video forgery detection based on non-subsampled contourlet transform and gradient information. Inf Technol J 11(10):1456–1462
20. Chen S, Tan S, Li B, Huang J (2016) Automatic detection of object-based forgery in advanced video. IEEE Trans Circ Sys Vid Tech 26(11):2138–2151
21. Chetty G, Biswas M, Singh R (2010) Digital video tamper detection based on multimodal fusion of residue features. In: 2010 Fourth international conference on network and system security. IEEE, pp 606–613
22. Chittapur GB, Murali S, Prabhakara H, Anami BS (2014) Exposing digital forgery in video by mean frame comparison techniques. In: Emerging research in electronics, computer science and technology. Springer, pp 557–562
23. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv:14061078
24. D'Amiano L, Cozzolino D, Poggi G, Verdoliva L (2018) A patch match-based dense-field algorithm for video copy-move detection and localization. IEEE Trans Circ Sys Vid Technol 29:669–682
25. Dataset (Accessed 2 Dec 2019) [Online] https://spectrum.ieee.org/tech-talk/artificial-intelligence/machine-learning/facebook-ai-launches-its-deepfake-detection-challenge
26. Dataset (Accessed 2 Nov 2019) Test Database: [Online] https://ceng2.ktu.edu.tr/gulutas/test_database.rar
27. D'Avino D, Cozzolino D, Poggi G, Verdoliva L (2017) Autoencoder with recurrent neural networks for video forgery detection. Electron Imag 2017(7):92–99
28. Dong Q, Yang G, Zhu N (2012) A mcea based passive forensics scheme for detecting frame-based video tampering. Digit Invest 9(2):151–159
29. Fadl SM, Han Q, Li Q (2018) Authentication of surveillance videos: detecting frame duplication based on residual frame. J Forensic Sci 63(4):1099–1109
30. Fan Y, Zhu YS, Liu Z (2016) An improved sift-based copy-move forgery detection method using t-linkage and multi-scale analysis. J Inf Hiding Multimedia Sign Process 7(2):399–408
31. Fayyaz MA, Anjum A, Ziauddin S, Khan A, Sarfaraz A (2020) An improved surveillance video forgery detection technique using sensor pattern noise and correlation of noise residues. Multimedia Tools Appl 79(9):5767–5788
32. Feng C, Xu Z, Zhang W, Xu Y (2014) Automatic location of frame deletion point for digital video forensics. In: Proceedings of the 2nd ACM workshop on information hiding and multimedia security, pp 171–179
33. Gironi A, Fontani M, Bianchi T, Piva A, Barni M (2014) A video forensic technique for detecting frame deletion and insertion. In: 2014 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 6226–6230
34. GRIP (2017) Splicing Dataset: [Online] http://www.grip.unina.it/download/prog/ForgedVideosDataset/Splicing/. Accessed 24 Nov 2019
35. GRIP (2018) Copy-move Dataset: [Online] http://www.grip.unina.it/download/prog/ForgedVideosDataset/Copymove/. Accessed 24 Nov 2019
36. Guan H, Kozak M, Robertson E, Lee Y, Yates AN, Delgado A, Zhou D, Kheyrkhah T, Smith J, Fiscus J (2019) Mfc datasets: Large-scale benchmark datasets for media forensic challenge evaluation. In: 2019 IEEE winter applications of computer vision workshops (WACVW). IEEE, pp 63–72
37. Güera D, Delp EJ (2018) Deepfake video detection using recurrent neural networks. In: 2018 15th IEEE international conference on advanced video and signal based surveillance (AVSS). IEEE, pp 1–6
38. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778

39. Hong JH, Yang Y, Oh BT (2019) Detection of frame deletion in hevc-coded video in the compressed domain. Digit Invest 30:23–31
40. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017) Mobilenets: efficient convolutional neural networks for mobile vision applications. arXiv:170404861
41. Hsu CC, Hung TY, Lin CW, Hsu CT (2008) Video forgery detection using correlation of noise residue. In: 2008 IEEE 10th workshop on multimedia signal processing. IEEE, pp 170–174
42. Hu X, Ni J, Pan R (2015) Detecting video forgery by estimating extrinsic camera parameters. In: International workshop on digital watermarking. Springer, pp 28–38
43. Hu Y, Li CT, Wang Y, Liu BB (2012) An improved fingerprinting algorithm for detection of video frame duplication forgery. Int J Digit Crime Forensics (IJDCF) 4(3):20–32
44. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4700–4708
45. Hyun DK, Lee MJ, Ryu SJ, Lee HY, Lee HK (2013) Forgery detection for surveillance video. In: The era of interactive media. Springer, pp 25–36
46. Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T (2014) Caffe: Convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM international conference on multimedia, pp 675–678
47. Johnston P, Elyan E (2019) A review of digital video tampering: from simple editing to full synthesis. Digit Invest 29:67–81
48. Johnston P, Elyan E, Jayne C (2019) Video tampering localisation using features learned from authentic content. Neural Comput Appl 32:12243–12257
49. Kancherla K, Mukkamala S (2012) Novel blind video forgery detection using markov models on motion residue. In: Asian conference on intelligent information and database systems. Springer, pp 308–315
50. Kang X, Liu J, Liu H, Wang ZJ (2016) Forensics and counter anti-forensics of video inter-frame forgery. Multimedia Tools Appl 75(21):13833–13853
51. Kharat J, Chougule S (2020) A passive blind forgery detection technique to identify frame duplication attack. Multimedia Tools Appl 79:8107–8123
52. Kingra S, Aggarwal N, Singh RD (2017) Inter-frame forgery detection in h. 264 videos using motion and brightness gradients. Multimedia Tools Appl 76(24):25767–25786
53. Kobayashi M, Okabe T, Sato Y (2009) Detecting video forgeries based on noise characteristics. In: Pacific-rim symposium on image and video technology. Springer, pp 306–317
54. Kobayashi M, Okabe T, Sato Y (2010) Detecting forgery from static-scene video based on inconsistency in noise level functions. IEEE Trans Inf Forensics Sec 5(4):883–892
55. Kono K, Yoshida T, Ohshiro S, Babaguchi N (2018) Passive video forgery detection considering spatio-temporal consistency. In: International conference on soft computing and pattern recognition. Springer, pp 381–391
56. Koopman M, Rodriguez AM, Geradts Z (2018) Detection of deepfake video manipulation. In: The 20th Irish machine vision and image processing conference (IMVIP), pp 133–136
57. Korshunov P, Marcel S (2018) Deepfakes: a new threat to face recognition? Assessment and detection. arXiv:181208685
58. KTH (Accessed 24 Nov 2019) Database:[Online] http://www.nada.kth.se/cvap/actions
59. Labartino D, Bianchi T, De Rosa A, Fontani M, Vázquez-Padín D, Piva A, Barni M (2013) Localization of forgeries in mpeg-2 video through gop size and dq analysis. In: 2013 IEEE 15th international workshop on multimedia signal processing (MMSP). IEEE, pp 494–499
60. Li F, Huang T (2014) Video copy-move forgery detection and localization based on structural similarity. In: Proceedings of the 3rd international conference on multimedia technology (ICMT 2013). Springer, pp 63–76
61. Li L, Wang X, Zhang W, Yang G, Hu G (2012) Detecting removed object from video with stationary background. In: International workshop on digital watermarking. Springer, pp 242–252
62. Li L, Li S, Zhu H, Chu SC, Roddick JF, Pan JS (2013) An efficient scheme for detecting copy-move forged images by local binary patterns. J Inf Hiding Multimedia Sig Process 4(1):46–56
63. Li Y, Lyu S (2019) Exposing deepfake videos by detecting face warping artifacts. In: The IEEE conference on computer vision and pattern recognition (CVPR) workshops
64. Li Y, Chang MC, Lyu S (2018) In ictu oculi: exposing ai created fake videos by detecting eye blinking. In: 2018 IEEE International Workshop on Information Forensics and Security (WIFS). IEEE, pp 1–7
65. Li Z, Zhang Z, Guo S, Wang J (2016) Video inter-frame forgery identification based on the consistency of quotient of mssim. Sec and Commun Netw 9(17):4548–4556. https://doi.org/10.1002/sec.1648
66. Liao SY, Huang TQ (2013) Video copy-move forgery detection and localization based on tamura texture features. 2013 6th International Congress on Image and Signal Processing (CISP) 2:864–868

67. Lin CS, Tsay JJ (2013) Passive approach for video forgery detection and localization. In: The second international conference on cyber security, cyber peacefare and digital forensic (CyberSec2013), pp 107–112
68. Lin CS, Tsay JJ (2014) A passive approach for effective detection and localization of region-level video forgery with spatio-temporal coherence analysis. Digit Invest 11(2):120–140
69. Lin GS, Chang JF (2012) Detection of frame duplication forgery in videos based on spatial and temporal analysis. Int J Pat Recogn Artif Intell 26(07):1250017
70. Liu H, Li S, Bian S (2014) Detecting frame deletion in h. 264 video. In: International conference on information security practice and experience. Springer, pp 262–270
71. Liu Y, Huang T (2017) Exposing video inter-frame forgery by zernike opponent chromaticity moments and coarseness analysis. Multimedia Sys 23(2):223–238
72. Liu Y, Huang T, Liu Y (2018) A novel video forgery detection algorithm for blue screen compositing based on 3-stage foreground analysis and tracking. Multimedia Tools Appl 77(6):7405–7427
73. Long C, Smith E, Basharat A, Hoogs A (2017) A c3d-based convolutional neural network for frame dropping detection in a single video shot. In: 2017 IEEE Conference on computer vision and pattern recognition workshops (CVPRW). IEEE, pp 1898–1906
74. Mathai M, Rajan D, Emmanuel S (2016) Video forgery detection and localization using normalized cross-correlation of moment features. In: 2016 IEEE Southwest symposium on image analysis and interpretation (SSIAI). IEEE, pp 149–152
75. Mizher MA, Ang MC, Mazhar AA, Mizher MA (2017) A review of video falsifying techniques and video forgery detection techniques. Int J Elec Sec Digit Forensics 9(3):191–208
76. Mondaini N, Caldelli R, Piva A, Barni M, Cappellini V (2007) Detection of malevolent changes in digital video for forensic applications. In: Security, steganography, and watermarking of multimedia contents IX, International Society for Optics and Photonics, vol 6505, p 65050T
77. Nguyen HH, Yamagishi J, Echizen I (2019) Capsule-forensics: using capsule networks to detect forged images and videos. In: ICASSP 2019-2019 IEEE International conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 2307–2311
78. Nguyen TT, Nguyen CM, Nguyen DT, Nguyen DT, Nahavandi S (2019) Deep learning for deepfakes creation and detection. arXiv:190911573
79. NIMBLE (Accessed 24 Nov 2019) Nimble Challenge 2017 Dataset: [Online] https://www.nist.gov/itl/iad/mig/nimble-challenge-2017-evaluation
80. NTHU (Accessed 22 Nov 2019a) Forensics Project Dataset [Online]: http://www.ee.nthu.edu.tw/cwlin/forensics/forensics.htm
81. NTHU (Accessed 22 Nov 2019b) Video Inpainting Project: http://www.ee.nthu.edu.tw/cwlin/inpainting/inpainting.htm
82. Oh S, Hoogs A, Perera A, Cuntoor N, Chen CC, Lee JT, Mukherjee S, Aggarwal J, Lee H, Davis L et al (2011) A large-scale benchmark dataset for event recognition in surveillance video. In: CVPR 2011. IEEE, pp 3153–3160
83. Pandey RC, Singh SK, Shukla K (2014) Passive copy-move forgery detection in videos. In: 2014 International conference on computer and communication technology (ICCCT). IEEE, pp 301–306
84. Pandey RC, Singh SK, Shukla KK (2016) Passive forensics in image and video using noise features: a review. Digit Invest 19:1–28
85. Pu H, Huang T, Guo G, Weng B, You L (2019) Video tampering detection algorithm based on spatial constraints and stable feature. In: UK workshop on computational intelligence. Springer, pp 541–553
86. Qadir G, Yahaya S, Ho ATS (2012) Surrey university library for forensic analysis (sulfa) of video content, pp 1–6. http://sulfa.cs.surrey.ac.uk/
87. Ravi H, Subramanyam AV, Gupta G, Kumar BA (2014) Compression noise based video forgery detection. In: 2014 IEEE international conference on image processing (ICIP). IEEE, pp 5352–5356
88. REWIND (2013) Datset: [Online]. https://sites.google.com/site/rewindpolimi/downloads/datasets/video-copy-move-forgeries-datase. Accessed 2 Nov 2019
89. Rocha A, Scheirer W, Boult T, Goldenstein S (2011) Vision of the unseen: current trends and challenges in digital image and video forensics. ACM Comput Surv (CSUR) 43(4):1–42
90. Rössler A, Cozzolino D, Verdoliva L, Riess C, Thies J, Nießner M (2018) Faceforensics: a large-scale video dataset for forgery detection in human faces. arXiv:180309179
91. Rossler A, Cozzolino D, Verdoliva L, Riess C, Thies J, Nießner M (2019) Faceforensics++: learning to detect manipulated facial images. In: Proceedings of the IEEE international conference on computer vision, pp 1–11
92. Sabir E, Cheng J, Jaiswal A, AbdAlmageed W, Masi I, Natarajan P (2019) Recurrent convolutional strategies for face manipulation detection in videos. Interfaces (GUI) 3:1

93. Sabour S, Frosst N, Hinton GE (2017) Dynamic routing between capsules. In: Advances in neural information processing systems, pp 3856–3866
94. Saddique M, Asghar K, Bajwa UI, Hussain M, Habib Z (2019) Spatial video forgery detection and localization using texture analysis of consecutive frames. Adv Elect Comput Eng 19(3):97–108
95. Shanableh T (2013) Detection of frame deletion for digital video forensics. Digit Invest 10(4):350–360
96. Shelke NA, Chatur P (2013) A survey on various digital video watermarking schemes. IJCSET 4(12)
97. Shelke NA, Chatur P (2016) Optimized and hybrid based watermarking system for digital video security. In: 2016 International conference on wireless communications, signal processing and networking (WiSPNET). IEEE, pp 1068–1074
98. Shih TK, Tang NC, Tsai JC, Hwang JN (2010) Video motion interpolation for special effect applications. IEEE Trans Syst Man Cybernetics Part C (Applications and Reviews) 41(5):720–732
99. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv:14091556
100. Singh G, Singh K (2019) Video frame and region duplication forgery detection based on correlation coefficient and coefficient of variation. Multimedia Tools Appl 78(9):11527–11562
101. Singh RD, Aggarwal N (2017a) Detection and localization of copy-paste forgeries in digital videos. Forensic Sci Int 281:75–91
102. Singh RD, Aggarwal N (2017b) Detection of upscale-crop and splicing for digital video authentication. Digit Invest 21:31–52
103. Singh RD, Aggarwal N (2017c) Optical flow and prediction residual based hybrid forensic system for inter-frame tampering detection. J Circ Syst Comput 26(07):1750107
104. Singh RD, Aggarwal N (2018) Video content authentication techniques: a comprehensive survey. Multimedia Syst 24(2):211–240
105. Singh VK, Pant P, Tripathi RC (2015) Detection of frame duplication type of forgery in digital video using sub-block based features. In: International conference on digital forensics and cyber crime. Springer, pp 29–38
106. Sitara K, Mehtre BM (2016) Digital video tampering detection: an overview of passive techniques. Digit Invest 18:8–22
107. Sitara K, Mehtre BM (2017) A comprehensive approach for exposing inter-frame video forgeries. In: 2017 IEEE 13th international colloquium on signal processing its applications (CSPA), pp 73–78. https://doi.org/10.1109/CSPA.2017.8064927
108. Stamm MC, Lin WS, Liu KJR (2012) Temporal forensics and anti-forensics for motion compensated video. IEEE Trans Inf Forensics Sec 7(4):1315–1329
109. Su L, Li C (2018) A novel passive forgery detection algorithm for video region duplication. Multidim Syst Sign Process 29(3):1173–1190
110. Su L, Huang T, Yang J (2015a) A video forgery detection algorithm based on compressive sensing. Multimedia Tools Appl 74(17):6641–6656
111. Su L, Li C, Lai Y, Yang J (2018) A fast forgery detection algorithm based on exponential-fourier moments for video region duplication. IEEE Trans Multimedia 20(4):825–840
112. Su L, Luo H, Wang S (2019) A novel forgery detection algorithm for video foreground removal. IEEE Access 7:109719–109728
113. Su PC, Suei PL, Chang MK, Lain J (2015b) Forensic and anti-forensic techniques for video shot editing in h. 264/avc. J Vis Commun Image Represent 29:103–113
114. Su Y, Zhang J, Liu J (2009) Exposing digital video forgery by detecting motion-compensated edge artifact. In: 2009 international conference on computational intelligence and software engineering. IEEE, pp 1–4
115. Subramanyam AV, Emmanuel S (2012) Video forgery detection using hog features and compression properties. In: 2012 IEEE 14th international workshop on multimedia signal processing (MMSP). IEEE, pp 89–94
116. Subramanyam AV, Emmanuel S (2013) Pixel estimation based video forgery detection. In: 2013 IEEE International conference on acoustics, speech and signal processing. IEEE, pp 3038–3042
117. SULFA (Accessed 1 Nov 2019) Surrey University Library for Forensic Analysis Dataset [Online]: http://sulfa.cs.surrey.ac.uk/
118. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2818–2826
119. Tan S, Chen S, Li B (2015) Gop based automatic detection of object-based forgery in advanced video. In: 2015 Asia-pacific signal and information processing association annual summit and conference (APSIPA). IEEE, pp 719–722

120. Tralic D, Grgic S, Zovko-Cihlar B (2014) Video frame copy-move forgery detection based on cellular automata and local binary patterns. In: 2014 X international symposium on telecommunications (BIHTEL). IEEE, pp 1–4
121. TREC (Accessed 2 Nov 2019) Video Retrieval Evaluation:: TRECVID Dataset [Online] http://trecvid.nist.gov/
122. Ulutas G, Ustubioglu B, Ulutas M, Nabiyev V (2017) Frame duplication/mirroring detection method with binary features. IET Image Process 11(5):333–342
123. Ulutas G, Ustubioglu B, Ulutas M, Nabiyev VV (2018) Frame duplication detection based on bow model. Multimedia Sys 24(5):549–567
124. VIUCM (Accessed 24 Nov 2019) Video Inpainting Under Camera Motion: [Online] http://www.tc.umn.edu/patw0007/video-inpainting/
125. VTD (Accessed 2 Nov 2019) Video Tampering Dataset: [Online]: https://www.youtube.com/channel/UCZuuu-iyZvPptbIUHT9tMrA
126. VTL (Accessed 24 Nov 2019) Video Trace Library: [Online] http://trace.eas.asu.edu/
127. Wahab AWA, Bagiwa MA, Idris MYI, Khan S, Razak Z, Ariffin MRK (2014) Passive video forgery detection techniques: a survey. In: 2014 10th international conference on information assurance and security. IEEE, pp 29–34
128. Wang Q, Li Z, Zhang Z, Ma Q (2014a) Video inter-frame forgery identification based on consistency of correlation coefficients of gray values. J Comput Commun 2(04):51
129. Wang Q, Li Z, Zhang Z, Ma Q (2014b) Video inter-frame forgery identification based on optical flow consistency. Sensors & Transducers 166(3):229
130. Wang W, Farid H (2006) Exposing digital forgeries in video by detecting double mpeg compression. In: Proceedings of the 8th workshop on multimedia and security, pp 37–47
131. Wang W, Farid H (2007) Exposing digital forgeries in interlaced and deinterlaced video. IEEE Trans Inf Forensics Sec 2(3):438–449
132. Wang W, Farid H (2007) Exposing digital forgeries in video by detecting duplication. In: Proceedings of the 9th workshop on multimedia & security, pp 35–42
133. Wang W, Farid H (2009) Exposing digital forgeries in video by detecting double quantization. In: Proceedings of the 11th ACM workshop on multimedia and security, pp 39–48
134. Wang W, Jiang X, Wang S, Wan M, Sun T (2013) Identifying video forgery process using optical flow. In: International workshop on digital watermarking. Springer, pp 244–257
135. Wang Y, Jodoin PM, Porikli F, Konrad J, Benezeth Y, Ishwar P (2014) Cdnet 2014: An expanded change detection benchmark dataset. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 387–394
136. Wei W, Fan X, Song H, Wang H (2019) Video tamper detection based on multi-scale mutual information. Multimedia Tools Appl 78(19):27109–27126
137. Wu Y, Jiang X, Sun T, Wang W (2014) Exposing video inter-frame forgery based on velocity field consistency. In: 2014 IEEE International conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 2674–2678
138. Xiph (Accessed 24 Nov 2019) Dataset [Online]: www.xiph.org
139. Xu J, Liang Y, Tian X, Xie A (2016) A novel video inter-frame forgery detection method based on histogram intersection. In: 2016 IEEE/CIC international conference on communications in China (ICCC). IEEE, pp 1–6
140. Yang J, Huang T, Su L (2016) Using similarity analysis to detect frame duplication forgery in videos. Multimedia Tools Appl 75(4):1793–1811
141. Yang X, Li Y, Lyu S (2019) Exposing deep fakes using inconsistent head poses. In: ICASSP 2019-2019 IEEE International conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 8261–8265
142. Yao H, Ni R, Zhao Y (2019) An approach to detect video frame deletion under anti-forensics. J Real-Time Image Proc 16(3):751–764
143. Yao Y, Shi Y, Weng S, Guan B (2017) Deep learning for detection of object-based forgery in advanced video. Symmetry 10(1):3
144. YFCC (Accessed 24 Nov 2019) 3YFCC100m Dataset: [Online] http://www.yfcc100m.org
145. Yin L, Bai Z, Yang R (2014) Video forgery detection based on nonnegative tensor factorization. In: 2014 4th IEEE international conference on information science and technology. IEEE, pp 148–151
146. Yu L, Wang H, Han Q, Niu X, Yiu SM, Fang J, Wang Z (2016) Exposing frame deletion by detecting abrupt changes in video streams. Neurocomputing 205:84–91
147. Zampoglou M, Markatopoulou F, Mercier G, Touska D, Apostolidis E, Papadopoulos S, Cozien R, Patras I, Mezaris V, Kompatsiaris I (2019) Detecting tampered videos with multimedia forensics and deep learning. In: International conference on multimedia modeling. Springer, pp 374–386

148. Zhang J, Su Y, Zhang M (2009) Exposing digital video forgery by ghost shadow artifact. In: Proceedings of the first ACM workshop on multimedia in forensics, pp 49–54
149. Zhang Z, Liu Q (2019) Detect video forgery by performing transfer learning on deep neural network. In: The international conference on natural computation, fuzzy systems and knowledge discovery. Springer, pp 415–422
150. Zhang Z, Hou J, Li Z, Li D (2015a) Inter-frame forgery detection for static-background video based on mvp consistency. In: International workshop on digital watermarking. Springer, pp 94–106
151. Zhang Z, Hou J, Ma Q, Li Z (2015b) Efficient video frame insertion and deletion detection based on inconsistency of correlations between local binary pattern coded frames. Sec Commun Netw 8(2):311–320
152. Zhao DN, Wang RK, Lu ZM (2018) Inter-frame passive-blind forgery detection for video shot based on similarity analysis. Multimedia Tools Appl
153. Zheng L, Sun T, Shi YQ (2014) Inter-frame video forgery detection based on block-wise brightness variance descriptor. In: International workshop on digital watermarking. Springer, pp 18–30