**Video Forgery Detection Methods- A Review**

**1. Introduction**

With the availability of large number of multimedia capturing and manipulation tools available today, there is a huge number of such multimedia files existing. But the authenticity of such contents must be examined before it can be used for various purposes like as evidence in digital forensics, or before publishing it through the various media. Video forgery detection methods proposed by researchers in this field try to address this issue. Some of the key terms related to this field are detailed below.

* 1. Digital Forensics

Digital forensics is a branch of forensic science encompassing the recovery and investigation of material found in digital devices, often in relation to computer crime.[[17]](https://en.wikipedia.org/wiki/Digital_forensics%22%20%5Cl%20%22cite_note-ijde-2002-1) Depending on the type of devices, media or artefacts, digital forensics investigation is branched into various types like computer forensics, mobile device forensics, network forensics, database forensics, forensic data analysis etc.

Computer forensics deals with analysing, recovering and presenting facts and opinions concerning the digital content. With the availability of large number of multimedia manipulation tools, the retrieved multimedia forensic data may not be trustworthy. Once computer forensics collects evidence of this type it must be examined to check its genuineness. So, audio and video forgery detection tools play their roles in digital forensics.

* 1. Video Forensics

Forensic video analysis is the scientific examination, comparison and/or evaluation of video in legal matters. The video forensics is carried out in a forensic lab equipped with tools where the analysis is done with utmost care and integrity. Police and judiciary make use of the video recording from surveillance cameras as evidence because they are easily available. A video forensic expert is responsible for interpreting the images and video recordings in order to extract and present the sequence of events which occurred and to compare and analyse the results. But in some situations, this becomes difficult due to the manipulations done on the video knowingly or unknowingly.

* 1. Video Forgery

Video Forgery is a technique of generating altered or fake videos by combining, altering or creating new video. Thus, the authenticity of such digital videos is questionable and needs to be verified.[18]. Based on the approaches taken for manipulation it can be categorized as

* + 1. Spatial forgery

In spatial forgery alterations are done within a frame of the video. It can be by repositioning of objects within the same frame, copying and placing objects in the frame in other regions in the same frame(copy-move) or by placing objects from other frames in the current frame(splicing).

* + - 1. Copy—move

A type of video forgery in which a frame region is copied and pasted within the same frame(spatial) or in another frame(temporal).

* + - 1. Splicing

Object taken from another video is pasted in the frames of a video.

* + 1. Temporal forgery

Alterations are done within the frames of the video. It includes copy-move, frame insertion/deletion, frame duplication etc.

* + - 1. Frame insertion/deletion

Frames from other video are inserted in the video, or frames from the same video may be duplicated. Also, some frames in the video can be deleted.

* + 1. Spatial temporal forgery

Alterations within the frame (intra frame) and in between the frames (inter frames) take place. So, it is called as spatial temporal forgery.

1. **Review of the related literature**

Several approaches have been taken by researchers to detect forgery in video. And still better methods are being discovered.

“An approach to detect video frame deletion under anti‑forensics”[1] states that the approaches to frame deletion detection can be categorized in to those based on the traces left after recompression after frame deletion and those based on the inter frame continuity. But some anti-forensic methods do invalidate several detection methods by making several modifications in the tampered video. So the methods developed must also consider the effects of anti-forensics in the tampered video. But according to the authors, there is no anti-forensic method which affects the inter frame continuity traces. The paper proposes an easy anti-forensic strategy to attack inter-frame continuity based forensic methods. Two frames on both sides of the FDP (frame deletion point) are taken as templates and uses interpolation to smooth the spike caused by frame deletion. And the paper also proposes a frame deletion detection method which overcomes the above proposed anti -forensic strategy. The frame residual in the actual frames and the interpolated frames are found to be near to zero in H.265/HEVC videos unlike in H.264/AVC videos where it is distinguishable. So the method makes use of the local and global, spatial and frequency domain features to detect frame deletion under anti-forensics.

Zampoglou et al.,2019 in their work “Detecting tampered videos with multimedia forensics and deep learning”[2] two forensic filters used for manual verification like those based on DCT coefficients and video re-quantisation errors, are combined with deep convolutional networks designed for image classification. The forgery detection methods are categorized in to double/multiple quantisation detection, inter-frame forgery detection, and region tampering detection. This paper deals with the third category, the region tampering detection. Parts of a video sequence are inserted in the frames of another video sequence. Here the assumption is that some invisible pattern originated from the capturing or compression process which is detectable is altered by the insertion of the foreign content.

The filters used produce visible output maps that can be analysed by humans viz. Q4 and Cobalt filter. Q4 filter is obtained using an NXN block in the video frame. Each such block transformed using DCT and the NXN coefficient arrays are created from blocks in the video frame, corresponding to each coefficient in the DCT of various blocks. Each array will be of size 1/N of the original image in each dimension. For getting arrays which are large enough the DCT of 2X2 blocks are taken and 3 of the arrays for coefficients (0, 1),(1,0) and (1,1) are displayed using RGB colour channels.

Cobalt filter makes use of the MPEG-4 re-quantization error to detect forgery. The video is re-quantized with a different constant quality level and calculate the per pixel value and an error video is created which depicts the differences with the original video. If the difference is high the intensity of the error video is high and the non-homogeneity in intensity of the error video shows difference in quantization parameters in the original frame itself which indicates frame manipulation.

Both the outputs which are RGB images are fed to Convolutional neural networks pre trained for classification.

 Vaishali et al.,2019[3] propose an algorithm which is based on the temporal difference between adjacent frames in the actual video and its reconstructed form. Video is reconstructed using the frame prediction error. Video tampering attacks are broadly classified in to spatial, temporal and spatio- temporal forgery. The method utilizes the intrinsic fingerprints left on the video after compression or video capture. The recompression of a MPEG video results in two distinct fingerprints-spatial and temporal. Spatial fingerprint occurs in a single I frame, when zero, or integer multiple of fixed GOP length frames are added or deleted. Temporal fingerprint occurs in the sequence of B frame or P frame prediction errors only if frames are added or deleted. When a frame is added or deleted from a fixed GOP, its structure gets altered and the prediction error will increase. Frame prediction error is used as a strong feature to detect forgery because no forger is able to delete these fingerprints from the video sequence. Its framework is as follows.

a. A video clip in MP4 or AVI format is taken as input and passed to proposed system for computation of prediction error vector and optical ﬂow.

b. The full length video sequence given as input is divided into frames and saved in a folder in form of JPEG images.

 c. The structural similarity is extended to measure similarity between two frames of video clip. Then the similarities between frames in the temporal domain are measured and used to calculate prediction error between two frames.

d. Frame prediction error for each frame is calculated and frames are reconstructed based on this error. The calculated errors are stored in Excel sheets for further use.

f. Frame prediction errors of these reconstructed frames are also calculated. And again stored in excel sheets.

g. These stored frame prediction errors are used to draw plots of actual video and its reconstructed form.

h. These plots easily show the variation between frame prediction errors of forged video and genuine video. So, by comparing these plots one can easily classify between forged and genuine video.

i. Finally, optical motion is calculated using method proposed by Lucas-Kanade for input video clip as well as predicted video clip. A plot is drawn to show optical ﬂow of each video clip. With the help of plot tampering can be accurately detected and localized. Already proven feature of optical ﬂow is used to verify correct classiﬁcation of video.

According to Mubbashar Saddique *et al.,2019*[5] spatial forgery results in inconsistency in the texture and micro patterns in the frames which can be found out from the difference in consecutive frames. For that purpose two features viz. Chrominance value of consecutive frame difference (CCD) and Discriminative Local Robust Binary Pattern (DRLBP) are made use of. Support Vector Machine (SVM) is used to find out forgery based on the feature vector which combines CCD and DRLBP. These descriptors incorporates both texture and shape information which makes it robust to noise, shape and side variations. The paper focuses on spatial tampering. When the difference of consecutive frames (DOCFs) is found out it clearly shows the traces of forgery in forged frames but not in authentic ones.

First the video is divided in to segments (VSs) of 30 frames each and the frames are extracted. The DOCFs are found out and the features are extracted which are then passed to the SVM model. The model returns a decision as whether the video is authentic or forged. A new descriptor called CCD-DRLBP is proposed to extract features which are efficient for forgery detection.

In “Toward video tampering exposure: inferring compression parameters from pixels” [8], the authors propose a method to detect QP of a H.264/AVC compressed video. Although other methods like the one by Boss et.al exists, it can only detect the QP of key frames only and the estimation of QP of individual patches in the frame shows much deviation although the averaged QP of a frame is somewhat accurate. This paper tries to resolve these issues and it finds QP for predicted frames as well. The model makes use of video and image datasets with labelled patches to find the QP. Three CNNs are pre-trained with different but not adjacent QP values. It is found that estimation becomes more accurate with large patch size and larger stride to minimize the correlation between patches. The performance is evaluated and improvements suggested. Paper concludes saying that the QP estimation of predicted frames are not as accurate as of key frames. Further improvements are recommended.

In “Video tampering localisation using features learned from authentic content” [9] it is found that vast majority of video compression is in H.264/AVI or MPEG2 format. The paper uses the similar methods as in [8] to estimate quantisation parameter, inter-intra frame type and frame delta directly from pixels. CNN is trained to identify these values. Frame delta is used to identify key frames. It locates tampering in some manipulated video by identifying distinct compression profiles in the same video.

In “Coarse-to-fine Copy-move Forgery Detection for Video Forensics” [10], the approaches to detect frame copy-move forgeries are categorized in to image feature based and video feature based. This paper proposes a video motion feature based approach to detect frame copy-move forgery. The methods already existing demands high computational cost, have unstable detection performance and limited applicability. The authors try to overcome these limitations by the proposed method.

The method depends on the Lucas-Kanade Optical Flow, proposed by B.D. Lucas and T. Kanade [15], to compute the OF for each frame. Merits of this method are rapid computation, simple application, and robustness under noise. OF between adjacent frames i and i+1 are found out and stored as OFi. Then the correlation between the OF of adjacent frames are calculated which ranges from -1 to +1. High correlation indicates higher similarity between frame pairs and hence genuineness.

To reduce the computational demand in finding the correlation to detect forgery, another feature OF sum is found out for each frame by adding the OXi and OYi of each frame for all pixel positions. Then in order to detect forgery OF sum consistency is made use of which either shows sudden spikes when frames are inserted in between or local symmetries when alterations in the video sequence are done carefully. The tampering location can be easily identified and further correlation calculations can be done once the location is identified greatly alleviating the computational burden.

Omar Ismael et al.,2015 in “Detection Of Video Forgery: A Review Of Literature “[11] state that video forgery primarily falls into two methods based on their approaches; active approaches and passive-blind approaches. In the present study, the performances of some typical video forgery algorithms are compared and an overview of passive digital video authentication method is demonstrated. It was noted that source identification methods hinges on the robust statistics features of source camera identification hardware such as Noise and CCD sensor features that are more dependable than camera software parts (CFA interpolation algorithms). Furthermore, it was found that quantifying double compression artefacts cause difficulty in localization of the forgery when the image is analysed or compressed via low quality factor in majority of methods. On the other hand, in some methods camera sensor noise was utilized to relate a distinct method-identification camera with a video and to determine tampering of video regions.

According to the authors, a statistical image model for splicing detection was proposed by Farid et.al , a version of which is the blind image forgery detection method that extracts features of classification via the Hilbert-Huang transform (HHT) and statistical model that hinges on the moments of characteristic functions. This involves the application of wavelets transform to differentiate the spliced region detection. Their findings showed that the method is able to detect high accuracy of passive splicing localization detection. Other authors [20] developed a method to detect suspicious regions in video recorded from a static scene with the help of noise characteristics of the acquisition device described through a noise level function (NLF) in frame sequence.

Some of the relevant works in the area are considered and evaluated.

“Video Forgery Detection Using HOG Features and Compression Properties” [12] say that the intrinsic properties of the video are used to detect copy-move tampering. The copy-move video forgery is classified in to spatial tampering and temporal tampering. There are methods proposed which are based on SIFT features matching, Fourier-Mellin Transform etc. which can be used for spatial copy-move forgery detection, but not for temporal forgery detection.

Video forgery detection based on noise characteristics may not work if the duplicated region belongs to the same video. SIFT features may not detect forgery for small patch size if a forged patch has undergone scaling transformation even if the feature is robust. Therefore, features such as HoG which are dense image or block wise descriptors are useful in detecting such form of tampering.

“A Frame Tampering Detection Algorithm for MPEG videos” is the work by Su et al.,2011[13] in which it is said that in MPEG-2 standard, the structure of group of pictures (GOP) defines the orientation of inter and intra frames in the temporal sequence. During tampering when some frames are inserted or removed the structure gets changed in the subsequent compression. The coding type change helps to detect tampering in MPEG videos.

In “Photo Forensics from JPEG Dimples” [14], Farid et al.,2017 state that the artefact is introduced in JPEG compression when the DCT coefficients after quantization are converted from floating point type to integer using ceiling, or floor mathematical operators rather than rounding operator. The artefact introduced by various camera models differs and this becomes helpful in forgery detection by analysing the associated correlation energy of different blocks.

Minyoung et al.,2018[16] propose a learning algorithm to detect visual image manipulations which trains the model using large dataset of real photographs. It is trained to check the self-consistency whether the image is created using the same image pipeline using the EXIF metadata of the image. It achieves acceptable performance even without seeing any manipulated image during training. A consistency classifier is learned for each EXIF tag using pairs of photographs and the learned model is used to estimate self-consistency given two input image patches. The method uses a Siamese network to check whether different pairs of patches in an image have the same value for all of the EXIF attributes. Then calculates the overall consistency by combining all the metadata attribute consistency values. A low consistency shows that the patches originated from two different sources.

Kelton A.P. da Costa et al.,2020 [21] discuss forgery detection using illumination-based texture descriptor and the FOA-SVNN based classifier for the datasets DSO-1, DSI-1 and gets an accuracy of 95.23% and 94.59% respectively.

 In “Recent advances in digital image manipulation detection techniques: A brief review”[22] the authors make it known that although many datasets have been released in the field of image manipulation detection the number of tampered images is very less. The generation of synthesized images will help to overcome the problem and enough data becomes available for training in neural network based methods. By combining many machine learning models for manipulation detection at multiple scales and by transfer learning a general purpose image manipulation detection system can be generated. The paper gives the details of the various publicly available datasets.

“An efficient approach for forgery detection in digital images using Hilbert–Huang transform”[23]deals with image forgery detection with post-processing attacks such as image compression, adding Gaussian noises or adjusting the contrast of the image and produce very high accuracies for the datasets CASIA –V1 ,CASIA-V2 , MICC-F2000 , MICC-F600 , MICC-F220 , CoMoFoD ,Internet websites and social media.

**3. Dataset**

The already existing datasets support any one type of forgery which can be used for evaluating such forgeries. But to evaluate and train for combination of spatial and temporal forgeries, new dataset may need to be created. The existing ones which seem useful are listed below.

SULFA forged [24] -contains 150 original videos of about 10 seconds duration and some spatio-temporal copy-move forgery videos for camera identiﬁcation.

TDTVD[27] dataset is created by removing events, objects, or persons at single or multiple locations in a video.

VTD[26]-33 tampered videos of 16 seconds duration. Contains three types of tampering- copy-move, splicing and swapping. Contains complete information about the tampering in the doctored videos.

InVID Fake Video Corpus[28] –contains 117 fake videos and 110 real videos with the annotations and descriptions.

 GRIP Dataset[25]- contains ten videos with splicing forgeries using Adobe After Effects.

The dataset for analysing the image forgeries include CASIA –V1 ,CASIA-V2 , MICC-F2000 , MICC-F600 , MICC-F220 , CoMoFoD, GRIP, Columbia Image Splicing dataset, Columbia Uncompressed etc.

**4. Steps Undertaken**

Which ever be the methodology adopted, the video forgery detection consists of the following steps.

**Step1: Feature extraction**

Features provide clues regarding the authenticity of frames. It may measure the discontinuity in spatial, temporal and frequency domain. It may also include both inter-frame and intra-frame forgery details, such that a single feature vector can handle both types of forgeries. Several techniques like Principal Component analysis(PCA), Discrete Cosine Transform(DCT), Scale Invariant Feature Transform(SIFT), Hilbert-Huang transform (HHT), Histogram of Oriented Gradients(HoG) and Difference of Consecutive Frames are used for feature extraction. Based on how much and what information is needed for the detection of video forgery, the required features are extracted from the video after the application of the necessary transformations like those mentioned above. Once the feature extraction process is completed the extracted features are used the next step.

**Step2: Classification**

For classification of the video as genuine or forged, there are many techniques available. The classification techniques can be categorized as

* Rule based
* Machine Learning based
* Deep Learning based

In rule based classification simple if then rules using some threshold may be used for classification. In Machine Learning(ML) based classifier machine learning models like SVM(Support Vector Machine), Decision Tree Classifier, Linear Regression or ANN(Artificial Neural Network) are used and in Deep Learning(DL) based classification, Deep Learning model like CNN, RNN etc are used for classification. Whether it is ML based or DL based, first the model has to be trained well using available information and only after that it can be used for the classification. Compared to ML based models, DL models require large volume of data for training the model. So, the choice of classifier very much depends on the dataset size. The type of available data and the type of the required result also affect the choice of the classifier. If all the fields in the data are not numeric, it must be converted to numeric format before supplying it for training and classification.

There are tools and libraries available in all popular programming languages which can be used easily to accomplish the task in both the above steps.

**5. Conclusion**

This paper presents a review of the works related to video forgery detection. Some of the techniques used are mentioned along with some publicly available datasets. Steps under taken in the process are explained which is applicable in any research project irrespective of the type of task. A clear flow of the research process is outlined in this paper.

**5.References**

[1] An approach to detect video frame deletion under anti‑forensics, Haichao Yao, Rongrong Ni, Yao Zhao ( 2019) Journal of Real-Time Image Processing

[2] Detecting tampered videos with multimedia forensics and deep learning Markos Zampoglou, Foteini Markatopoulou, Gregoire Mercier, Despoina Touska, Evlampios Apostolidis, Symeon Papadopoulos, Roger Cozien, Ioannis Patras, Vasileios Mezaris, and Ioannis Kompatsiaris Proc. 25th Int. Conf. on Multimedia Modeling (MMM 2019), Springer

[3] Tampering detection and localization in digital video using temporal difference between adjacent frames of actual and reconstructed video clip,Vaishali JoshiSanjay Jain,Bharati Vidyapeeth’s Institute of Computer Applications and Management 2019 Int. j. inf. tecnol.

[5.] Spatial video forgery detection and localization using texuture analysis of consecutive frames , Mubbashar SADDIQUE *et al.,(2019)*Advances in the electrical and computer engineering

[6]. Pevny, T., Bas, P., Fridrich, J.: Steganalysis by subtractive pixel adjacency matrix. IEEE Trans. Inf. Forensics Secur. (2010)

[8] JOHNSTON, P., ELYAN, E. and JAYNE, C. 2018. Toward video tampering exposure: Inferring compression parameters from pixels ,19th International Conference on Engineering Applications of Neural Networks (EAN). Springer International Publishing

 [9] Pamela Johnston, Eyad Elyan, and Chrisina Jayne. Video tampering localisation using features learned from authentic content, Neural Computing and Applications, Springer, 2019

[10] Coarse-to-fine Copy-move Forgery Detection for Video Forensics SHAN JIA,ZHENGQUAN XU, HAO WANG, CHUNHUI FENG, AND TAO WANG, CVPR Workshop 2019 IEEE Access

[11] DETECTION OF VIDEO FORGERY: A REVIEW OF LITERATURE, OMAR ISMAEL AL-SANJARY1, GHAZALI SULONG (2015)Journal of Theoretical and Applied Information Technology. Vol.74 No.2

[12] VIDEO FORGERY DETECTION USING HOG FEATURES AND COMPRESSION PROPERTIES, A.V. Subramanyam, Sabu Emmanuel 2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP) 2012 IEEE

[13] A Frame Tampering Detection Algorithm for MPEG videos ,Yuting Su, Weizhi Nie, Chengqian Zhang, 6th IEEE Joint International Information Technology and Artificial Intelligence Conference 2011 IEEE

[14] Photo Forensics from JPEG Dimples, Shruti Agarwal and Hany Farid 2017 IEEE Workshop on Information Forensics and Security (WIFS)

[15] B. D. Lucas, and T. Kanade, "An iterative image registration technique with an application to stereo vision." International Joint Conference on Artificial Intelligence. Morgan Kaufmann Publishers Inc 1981.

[16] Fighting Fake News: Image Splice Detection via Learned Self-Consistency Minyoung Huh Andrew Liu Andrew Owens Alexei A. Efros,ECCV2018

 [17]M Reith; C Carr; G Gunsch (2002). "An examination of digital forensic models". *International Journal of Digital Evidence.*

*[18]* A Review of Video Forgery and Its Detection Rohini Sawant and Manoj Sabnis IOSR Journal of Computer Engineering (IOSR-JCE) (2018)

[20] Kobayashi, M., Okabe, T., & Sato, Y. (2010). Detecting forgery from static-scene video based on inconsistency in noise level functions, *Information Forensics and Security*

*[21]* Kelton A.P. da Costa et.al (2020) A critical literature survey and prospects on tampering and anomaly detection in image data, Applied Soft Computing Journal

[22] Rahul Thakur, Rajesh Rohilla (2020) Review Article Recent advances in digital image manipulation detection techniques: A brief review, *Forensic Science International*

*[23]* H. Kasban a,∗ , Sabry Nassar b(2020) “An efficient approach for forgery detection in digital images using Hilbert–Huang transform”, Applied Soft Computing Journal

[24] Surrey University Library for Forensic Analysis (SULFA) of video content January 2012 DOI:  [10.1049/cp.2012.0422](http://dx.doi.org/10.1049/cp.2012.0422%22%20%5Ct%20%22https%3A//www.researchgate.net/publication/261046574_Surrey_University_Library_for_Forensic_Analysis_SULFA_of_video_content/_blank) Conference: Image Processing (IPR 2012)

*[25]* D’Avino D, Cozzolino D, Poggi G, Verdoliva L (2017) Autoencoder with recurrent neural networksfor video forgery detection. Electron Imaging 2017(7):92–99. https://doi.org/10.2352/ISSN.2470-1173.2017.7.MWSF-330

*[26]* Al-Sanjary OI, Ahmed AA, Sulong G (2016) Development of a video tampering dataset for forensicinvestigation. Forensic Sci Int 266:565–572. https://doi.org/10.1016/j.forsciint.2016.07.013

*[27]* Video tampering dataset development in temporal domain for video forgery authentication September 2020 Multimedia Tools and Applications 79(33-34) DOI: 10.1007/s11042-020-09205-w Hitesh D. Panchal,Dr. Hitesh Shah

*[28]* InVID Fake Video Corpus 2018 (v2.0), O. Papadopoulou, S. Papadopoulos, M. Zampoglou, I. Kompatsiaris (CERTH-ITI), D. Teyssou (AFP) https://mklab.iti.gr/results/fake-video-corpus/H. Kasban a,∗ , Sabry Nassar b(2020) “An efficient approach for forgery detection in digital images using Hilbert–Huang transform”, Applied Soft Computing Journal