ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 1, pp: (1-9), Month: April 2022 – September 2022, Available at: www.paperpublications.org

Video Forgery Detection: State-of-The-Art Review

[1*]Xuan Hau Nguyen, [2]Thai Son Tran

[1*]Corresponding Author: Xuan Hau Nguyen, Mientrung University of Civil Engineering, VietNam
[2]Mientrung University of Civil Engineering, VietNam

DOI: https://doi.org/10.5281/zenodo.6536002

Published Date: 10-May-2022

Abstract: **Nowadays, smartphones, camcorders, and security cameras are extensively used in many areas of daily life, such as offices, traffic lights, houses, dormitories, and more. Besides that, video content editing software like Window Movie Maker, Video Editor, Adobe Photoshop, Adobe After Effect, etc., are also widely available. They have many methods for editing video content easily. So, anyone can edit video content at their willing; even edited content contrast with original content. Recently, the rapid development of deep learning-based techniques have created deepfake videos with characters' faces replaced by other faces automatically, such as FakeApp, Faceswap, etc. That leads to "seeing is no longer believing.". Also, an authentic video provides stronger evidence in court. Therefore, video forgery detection proves that video authenticity has become an urgent requirement today.**

Keywords: **Video forgery detection; Video forensic; Deepfake video detection; Deep learning; convolutional neural network.**

1. INTRODUCTION

Video forgeries are classified into two main categories: video forgeries manipulated by humans and deepfake videos automatically manipulated by software tools.

1.1 Detecting Video Forgeries

The methods detect video forgeries manipulated by Humans [1] divided into passive and active methods. Passive methods only analyze video content to find traces of forgery. Meanwhile, active methods use given information such as Watermarking or Signature, which is inserted into the video, then that information is checked. If it doesn't change, then the video is authentic. Currently, Most videos do not usually insert given information, so passive methods have become a hot topic of interest to researchers. The passive methods can be classified into two categories: Video Inter-Frame Forgery and Region Tampering. Video Inter-Frame Forgery detects manipulated frames such as Frame Deletion, Frame Duplication, Frame Insertion, and Frame Shuffling. Meanwhile, the Video Intra-Frame Forgery methods detect manipulated small areas inside frames like Area Insertion, Area Deletion, and Area Duplication.

Recently with the steady development of video editing applications, 3-D regions in videos are copied, then pasted to other positions, and edited the brightness, geometry, and similar things have been edited easily. The 3-D regions can be small 3-D regions inside consecutive frame sequences or entire consecutive frame sequences [2, 3]. Those forgeries have become a popular tampering method used in video tampering, and it isn't very easy to detect naked eyes. So, it has been urging to find algorithms to detect this kind of video forgery. Besides that, manipulations at frame level conceal or imitate objects in the video. These manipulations are simple skills in editing the content of the video. But, they would create forged videos that are hard to detect, especially with naked eyes. In addition, manipulations of forged videos at the frame level were strongly supported by video content editing applications such as Movie Maker, Photoshop, or After Effect. Anyone can perform duplication, deletion, or insertion of a-frames sequence by one or two actions on these applications.

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 1, pp: (1-9), Month: April 2022 – September 2022, Available at: www.paperpublications.org

Nowadays, there are many methods to detect forgery videos. Most of them are based on handicraft features analysis of video frames [4]. These features are optical flow, color histogram, texture, noise, motion energy, correlation coefficients of grey values, and singular value decomposition (SVD). Analysis of these features is on a large number of frames in a video, it usually consumes a lot of time, and it has become a significant challenge. Through recent researches [5], deep learning has outstanding results. Notably, the convolutional neural networks (CNNs) have achieved excellent results in solving many challenging vision problems such as self-driving cars, visual captioning, object detection, and especially in large-scale image recognition. That has motivated many researchers who have applied CNN models to detect video forgeries.

1.2 Detecting Deepfake Videos

Besides manipulated video forgery by humans, millions of videos are recently uploaded on the internet every day; many videos in them have been manipulated by fully automated techniques change video content. And the development of that techniques has raised dangerous consequences for society and individuals. Especially in the last three years, Deep learning-based face replacement techniques in the video have been rapidly developed; these crucial tools are Faceswap [1], Faceswap-GAN [6], DeepFaceLab [3] and DFaker [4], etc. used to create videos in which contain face tampering. The naked eye hardly distinguishes these facial video forgeries. They can be made for malicious purposes such as pornographic videos of celebrities, politicians, fake news, policy tensions, and fake surveillance videos. So, currently, facial video forgery detection has become a hot topic of interest amongst researchers in the world. In addition, in recent years, Facebook, Microsoft, some largest universities, and partners have organized the Deepfake Detection Challenges that are designed to incentivize rapid progress in this area by inviting participants to compete to create new methods to detect and prevent modified media content.

Deepfake was defined by merging deep learning and fake, one of the most powerful techniques to create forged multimedia content. Deepfake videos were facial faked videos by the face of the target person that is transferred to a video that has a source person to create a new video that has the face of the target person and actions as the source person. Besides the fast development of hardware, Deepfake models based on autoencoders [7] or generative adversarial networks (GAN) [8], which have created realistic fake videos by automatically. It eliminates manual editing steps. For example, Faceswap, Faceswap-GAN tools, etc., which can be used to produce a realistic fake video by anyone who is only basically trained with most of the action via Graphical User Interface with no effort. Deepfake videos are malicious when they are fake politician videos or fake celebrity pornography videos. It is not only an awful influence on an individual but also on the community [5].

Recently, there have been some suggested methods for detecting facial tampered video; most of them are only based on steganalysis features or learned features on spatial or temporal separately. Features that have relation in spatial and temporal in the video are not exploited. Because a video is a set of consecutive frames in temporal, all of the current methods have not given good results, and it is still a significant challenge.

2. BACKGROUND

With the explosion of information technology, uploading, downloading, creating, and watching videos online is a daily affair of people in today's society. Fake videos have had enormous consequences for individuals and society, so video authenticity is essential. In order to learn and understand the forensic video methods, the following knowledge is required to grasp the following.

2.1 Video Forgeries by Humans

Video Forgery by Humans generates fake videos by combining, altering, or creating a new video from the original videos. That can be classified into two categories: Video Inter-Frame Forgery and Video Intra-Frame Forgery. The Video Inter-Frame Forgery would create forged videos by modifying the whole frame; meanwhile, the Video Intra-Frame Forgery changing small areas inside the frame.

Figure 1 and Figure 2 are examples of Video Intra-Frame Forgery. Fig. 1 shows examples of video forgery by the copy-move method. The red areas were copied from other positions in the same video, and then they were pasted into the video to duplicate or hide objects in the same video. In figure 1(a), the action in the red area was duplicated. In Figure 1(b), the car in the red area was hidden. Figure 2 is examples of video forgery by splicing method. The red areas were copied from

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 1, pp: (1-9), Month: April 2022 – September 2022, Available at: www.paperpublications.org

another video, and then they were pasted into the video to insert or hide objects in the video. In figure 2(a), the face in the red area is copied from another video, and in Figure 2(b), the car plate in the red area was replaced by the other one.

In Figure 3 are examples of Video Inter-Frame Forgery. In Figure 3(a), the whole frames in red areas were copied from another position in the same video to duplicate the run man. In Figure 3(b), the entire frames in red areas were copied from another location in the same video to hide the man. The literature review of these video forgery detection methods would be presented in 3.1 section.

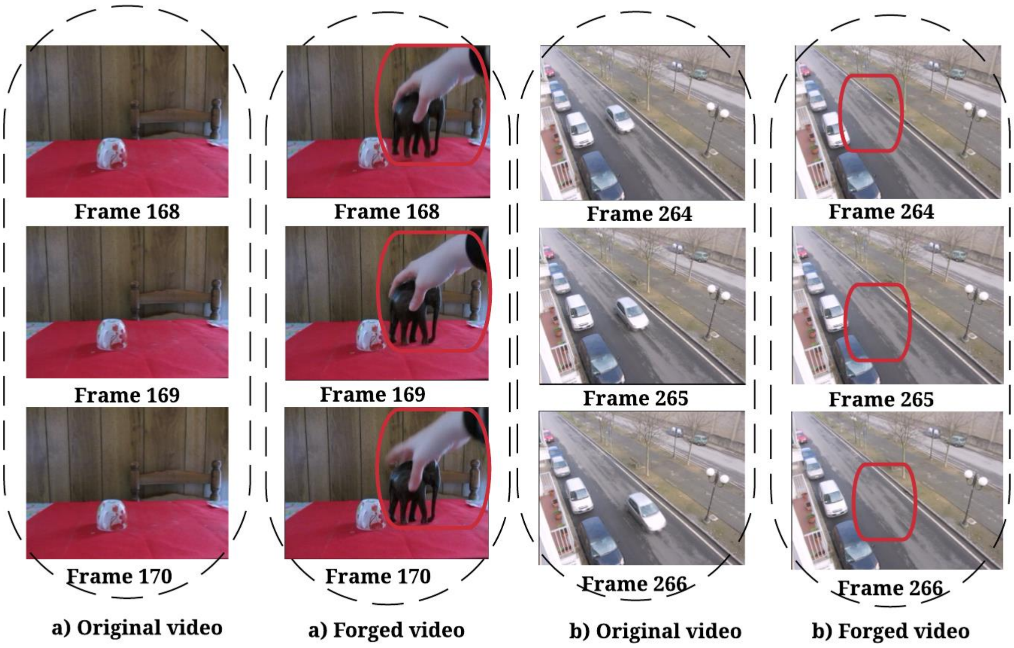

Fig. 1. Examples of video forgery by copy-move method. The red areas were copied from other positions in the same video, and then they were pasted to duplicate or hide objects in the same video.

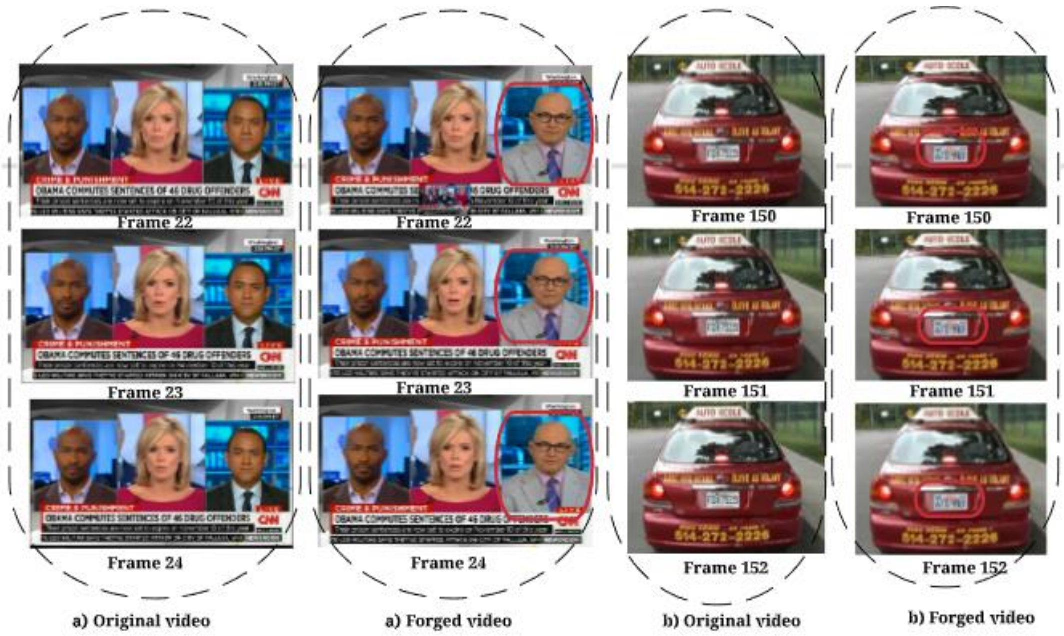

Fig. 2. Examples of video forgery by splicing method. The red areas were copied from other videos, and then they were pasted into the video to create the forged video.

ISSN 2350-1022

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 1, pp: (1-9), Month: April 2022 – September 2022, Available at: www.paperpublications.org

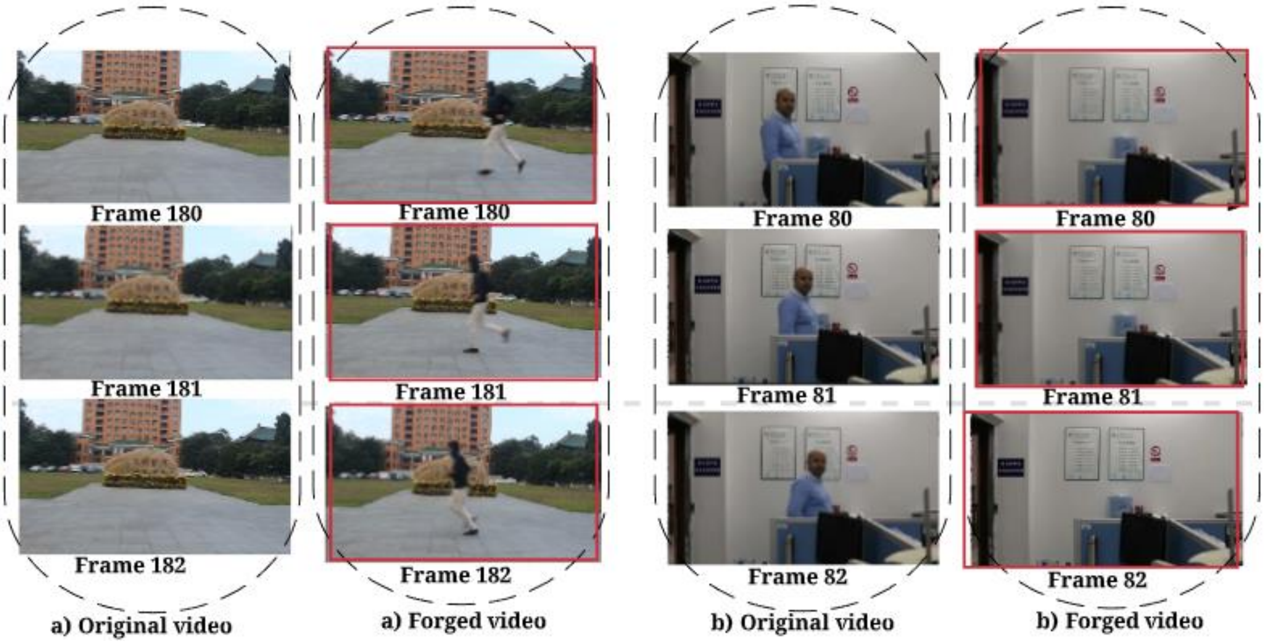

Fig. 3. Examples of video inter-frame forgery by copy-move method. The whole frames in red areas were copied from other positions in the same video, and then they were pasted to duplicate or hide objects in the same video.

2.2. Video Forgeries by Automatic Technique

An outstanding automated video forgery technique today is the Deepfake technique. Deepfake is derived from "deep learning" and "fake". Figure 4 is examples of Deepfake images cut from Deepfake videos. The photos numbered 1 are the original images, while the images numbered 2 are the Deepfake images. This technique can replace face images of a target person with the face of a source person to create videos that have the target person doing or saying things the source person does. Deepfake algorithms have used Deep learning models such as autoencoders and GANs that normally require a huge amount of image and video data to train models to create photo-realistic images and videos. Because politicians and celebrities may have a lot of images and videos available on the internet, so they are the top target of Deepfake. The literature review of Deepfake video detection methods would be presented in the 3.2 section.

Fig. 4: Examples of Deepfake images, the photos numbered 1 are the original images, while the images numbered 2 are the Deepfake images.

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 1, pp: (1-9), Month: April 2022 – September 2022, Available at: www.paperpublications.org

3. LITERATURE REVIEW

This section would describe the typical methods of detecting video forgery by humans as well as methods of detecting Deepfake videos.

3.1 Detecting Video Forgeries

Video forensic methods are divided into passive and active methods. Passive methods only analyze video content to detect traces of forgery. Meanwhile, Active methods use given information such as Signature or Watermarking, which is put in the video, then this information is checked. If it does not alter, that video is authentic otherwise forged. Nowadays, most videos don't usually insert given information, so passive methods have become a hot topic attracting researchers. Passive methods can be classified into two categories as Video Intra-Frame Forgery and Video Inter-Frame Forgery.

The Video forensic methods of Video Inter-Frame Forgery detect manipulated frames such as Frame Insertion, Frame Deletion, Frame Duplication, and Frame Shuffling. Meanwhile, the Video Intra-Frame Forgery methods detect manipulated small areas inside frames like Area Insertion, Area Deletion, and Area Duplication.

3.1.1 The Video forensic Methods of Video Inter-Frame Forgery

Some typical forensic video methods of video inter-frame forgery are as follows: Wang and Farid [9] was the first. The authors have described two techniques for detecting forgery in MPEG video sequences. Both of those techniques exploit the fact that static and temporal artifacts are introduced when a video sequence is subjected to double MPEG compression. These two techniques are very sensitive to noise and when changing frames in the same Group of Pictures (GOP). Because a video is often saved in compression after recording, so a forged video is compressed again after editing. There are some methods that depend on the codec of video, such as [10] used statistic pattern in the distribution of discrete cosine transform (DCT) coefficients to exploit double compression of the frame I. In [11], the authors calculated the pixel value in GOP. If there is any difference between the actual and calculated values, the video may be compressed many times. In [12], Markov statistics of compression, noise is used to detect double compression, and in [13], with assuming that the second compression has the same parameters as the first time of compression at which most of the methods fail, and the basic idea is when a frame is recompressed with the same quantization matrix again. The number of different DCT coefficients between the sequential two versions will decrease monotonically.

Besides that, some methods have depended on the correlation value of features on frames. As examples, in [14], the authors used differences in correlation coefficients of grey values between sequential frames to detect frame deletion and frame insertion. In [15], the authors calculated the spatial and temporal correlation to detect duplication of frame sequences. In [16], they used the correlation of SVD features of frame sequences to detect duplication of frame sequences. In [17], the authors used the correlation coefficient of sequences of frames DCT means. In [18, 19], used optical flow to detect frame insertion and frame deletion. In [20], the authors used Zernike opponent chromaticity moments to detect frame insertion, frame deletion, and duplication of frame sequences. And in [21], the first time the detection of a frame sequence duplication based on a deep convolutional neural network, this approach is most closely related to [22] that used the available model CNN to detect video inter-frame forgeries. This method used the I3D model [23] to detect duplicated frame sequences, which has high computational complexity. So, this method is not suitable for big data.

3.1.2 The Video Forensic Methods of Video Intra-Frame Forgery

The Video Forensic Methods of Video Intra-Frame Forgery exploit abnormal pixel values in a video to detect video forgery. Typical methods are as follows, Wang and Farid in [24] the authors presented two techniques for detecting tampering in deinterlaced and interlaced video. For deinterlaced video, authors explicitly model the correlations introduced by deinterlacing algorithms and show how tampering can destroy these correlations. For the interlaced video, the authors measure the inter-field and inter-frame motions that an authentic video is the same, but for a doctored video may be different. In both cases, they can localize tampering both in time and in space (what part of a video frame was manipulated). Compression artifacts make it somewhat more difficult to estimate the deinterlacing correlations, so these approaches are most appropriate for high-quality video. Wang and Farid [15] using the correlation between two frames of a video. This method is tampering detection in moving cameras. It would fail on a surveillance camera. Correlation of noise residue was used in [25]. This method is sensitive to quantization noise. Chen et al. [26] used features extracted from the motion residuals of each frame, and this method can not extract the localization of forged objects in the video frame. In [27], authors used ghost shadow artifacts to detect deletion of moving objects by video inpainting and used ghost shadow artifacts were gotten from the inconsistencies of moving foreground, segmented from the video frames, and the moving track obtained from the accumulative frame differences. This method is also sensitive to noise.

In [28], it calculated motion vector from adjacent frames for detecting object deletion in videos; this method may fail when tampering video is done carefully. In [29], it used temporal noise to detect two-dimensional regions copy-move; this method's accuracy depends on post-processing in forgery. In [30-32], the authors used features (Histogram of Oriented Gradients - HOG, Scale Invariant Feature Transform - SIFT or K-Singular Value Decomposition) to detect two-dimensional regions copy-move. These methods are sensitive when there are many similar objects in the video. In [33], residual of consecutive frames and in [34] Zernike moments and 3-D patch match are used to detect a two or three-dimensional region and paste it many times. These methods have low accuracy. In [35] Lucas Kanade, optical flow between adjacent frames is used to detect tampering copy-move regions; this method gives a low accuracy with low-quality videos.

3.2 Detecting Deepfake Video

Recently, Deepfake videos have had a strong negative impact on privacy, social security, and democracy [36]. Therefore, some methods for detecting Deepfake videos have been proposed. They can be divided into two groups: methods based on temporal features and methods based on spatial features.

3.2.1 Temporal features based deep learning methods

Most of the techniques in this group have been based on the observation that an original video is a sequence of consecutive frames and that the pixel values in the video have a significant spatial-temporal consistency. Otherwise, a Deepfake video is created from processing facial synthesis by a frame privately. So, the values of pixels in the synthesis regions are not consistent in Spatio-temporal. There are typical proposed methods as follows: Guera et al. [38] showed that there are intra-inconsistencies and temporal inconsistencies between frames. The authors combined CNN to extract features from frames and the long short-term memory (LSTM) network to capture temporal inconsistencies between frames. Likewise, Sabir et al. [37] used a convolutional neural network to extract features from the face area of frames. Then these features were inputted into the recurrent convolutional network (RNN) to exploit temporal information from consecutive frames. On the other hand, Li et al. [39] used eye blinking signals to distinguish between Deepfake videos and original ones based on the observation that in a Deepfake video, a person's eye is less frequent blinking than a person's that in an original video. The bounding boxes of eye landmark point sequences were created after a few pre-processing steps, such as aligning face, extracting, and scaling. They were inputted into the long-term recurrent convolutional networks for distinguishing eye blinking sequences between tampered videos and original videos. Most of the methods in this group were based on deep recurrent networks that learn temporal patterns across video frames.

3.2.2 Spatial features based deep learning methods

The difference with temporal features-based methods, methods in this group explore artifact patterns inside single frames separately in videos. These methods were based on the observation that Deepfake videos usually were created by an affine face wrapping approach to match original faces in the source videos on the frame privacy, which creates inconsistency between the warped face area and the surrounding context. There are typical proposed methods as follows: Li et al. [40] used the power models in deep learning such as VGG and ResNet to capture those inconsistent features at frames privacy in Deepfake videos. Similarly, in [41], the authors proposed two deep learning models with a low number of layers to distinguish frames between Deepfake videos and original videos. This approach has experimented with a dataset with high compressed videos, but the result is not so good. In [42], the authors applied a capsule network to get features that output from the VGG-19 network to distinguish the Deepfake videos from the original ones. In [43], the authors proposed 3D CNN models for detecting deepfake videos.

3.2.3 Spatial feature-based shallow classifiers

Some methods used a physiological signal for classifying Deepfake and original faces. For instance, in [44], the authors estimated 3-D head poses from the face area, which is based on 68 facial landmarks of the central face region, then an SVM classifier is used to classify Deepfake and original faces. Similarly, in [45], the authors used features such as eyes, teeth, and facial contours for detecting Deepfake videos. Besides that, the photoresponse non-uniformity (PRNU) was also used in [46] to distinguish the Deepfake videos from the original ones.

4. CONCLUSION

With the explosion of camera use in all fields, surveillance camera systems are equipped everywhere, and almost every smartphone user has a camera. So creating videos from cameras is normal, and anyone can create videos and then publish them on the internet. Besides, video content editing tools are more and more plentiful, convenient, user friendly, and have many more powerful features. Even-untrained users can also edit video content easily. In addition, recently, automatic content editing tools such as Deepfake technique have been developed strongly. People with a basic level of information technology can also use them to edit video content automatically, such as changing the face of a character in a video by another person's face. The content of the video changes may distort the truth, even contrary to the content of the original video. That is dangerous for both individuals and societies. Therefore that is the motivation for promoting research to detect video forgeries.

REFERENCES

[1] Milani, S., et al., An overview on video forensics. APSIPA Transactions on Signal and Information Processing, 2012. 1.

[2] Bestagini, P., et al., REWIND Video: copy -move forgeries dataset.https://sites.google.com/site/rewindpolimi/downloads/datasets/video-copy-move-forgeries-dataset. 2012.

[3] Qadir, G., S. Yahaya, and A.T. Ho, Surrey university library for forensic analysis (SULFA) of video content. 2012.

[4] Sitara, K. and B.M. Mehtre, Digital video tampering detection: An overview of passive techniques. Digital Investigation, 2016. 18: p. 8-22.

[5] Fan, J., C. Ma, and Y. Zhong, A selective overview of deep learning. arXiv preprint arXiv:1904.05526, 2019.

[6] 2.2.0, F., https://www.malavida.com/en/soft/fakeapp/#gref.

[7] Vincent, P., et al., Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. Journal of machine learning research, 2010. 11(12).

[8] Goodfellow, I., et al. Generative adversarial nets. in Advances in neural information processing systems. 2014.

[9] Wang, W. and H. Farid. Exposing digital forgeries in video by detecting double MPEG compression. in Proceedings of the 8th workshop on Multimedia and security. 2006.

[10] Wang, W. and H. Farid. Exposing digital forgeries in video by detecting double quantization. in Proceedings of the 11th ACM workshop on Multimedia and security. 2009. ACM.

[11] Subramanyam, A. and S. Emmanuel. Pixel estimation based video forgery detection. in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. 2013. IEEE.

[12] Ravi, H., et al. Compression noise based video forgery detection. in 2014 IEEE International Conference on Image Processing (ICIP). 2014. IEEE.

[13] Huang, Z., F. Huang, and J. Huang. Detection of double compression with the same bit rate in MPEG-2 videos. in 2014 IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP). 2014. IEEE.

[14] Wang, Q., et al., Video inter-frame forgery identification based on consistency of correlation coefficients of gray values. Journal of Computer and Communications, 2014. 2(04): p. 51.

[15] Wang, W. and H. Farid. Exposing digital forgeries in video by detecting duplication. in Proceedings of the 9th workshop on Multimedia & security. 2007. ACM.

[16] Yang, J., T. Huang, and L. Su, Using similarity analysis to detect frame duplication forgery in videos. Multimedia Tools and Applications, 2016. 75(4): p. 1793-1811.

[17] Singh, G. and K. Singh, Video frame and region duplication forgery detection based on correlation coefficient and coefficient of variation. Multimedia Tools and Applications, 2018: p. 1-36.

[18] Chao, J., X. Jiang, and T. Sun. A novel video inter-frame forgery model detection scheme based on optical flow consistency. in International Workshop on Digital Watermarking. 2012. Springer.

[19] Jia, S., et al., Coarse-to-fine copy-move forgery detection for video forensics. IEEE Access, 2018. 6: p. 25323-25335.

[20] Liu, Y. and T. Huang, Exposing video inter-frame forgery by Zernike opponent chromaticity moments and coarseness analysis. Multimedia Systems, 2017. 23(2): p. 223-238.

[21] Long, C., A. Basharat, and A. Hoogs, A Coarse-to-fine Deep Convolutional Neural Network Framework for Frame Duplication Detection and Localization in Video Forgery. arXiv preprint arXiv:1811.10762, 2018.

[22] Nguyen, X.H., et al., Detecting Video Inter-Frame Forgeries Based on Convolutional Neural Network Model. International Journal of Image, Graphics and Signal Processing, 2020. 12(2): p. 1.

[23] Carreira, J. and A. Zisserman. Quo vadis, action recognition? a new model and the kinetics dataset. in proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.

[24] Wang, W. and H. Farid, Exposing digital forgeries in interlaced and deinterlaced video. IEEE Transactions on Information Forensics and Security, 2007. 2(3): p. 438-449.

[25] Hsu, C.-C., et al. Video forgery detection using correlation of noise residue. in 2008 IEEE 10th workshop on multimedia signal processing. 2008. IEEE.

[26] Chen, S., et al., Automatic detection of object-based forgery in advanced video. IEEE Transactions on Circuits and Systems for Video Technology, 2015. 26(11): p. 2138-2151.

[27] Zhang, J., Y. Su, and M. Zhang. Exposing digital video forgery by ghost shadow artifact. in Proceedings of the First ACM workshop on Multimedia in forensics. 2009.

[28] Li, L., et al. Detecting removed object from video with stationary background. in International Workshop on Digital Watermarking. 2012. Springer.

[29] Kobayashi, M., T. Okabe, and Y. Sato, Detecting forgery from static-scene video based on inconsistency in noise level functions. IEEE Transactions on Information Forensics and Security, 2010. 5(4): p. 883-892.

[30] Pandey, R.C., S.K. Singh, and K. Shukla. Passive copy-move forgery detection in videos. in 2014 International Conference on Computer and Communication Technology (ICCCT). 2014. IEEE.

[31] Subramanyam, A. and S. Emmanuel. Video forgery detection using HOG features and compression properties. in 2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP). 2012. IEEE.

[32] Su, P.C., et al., Forensic and anti-forensic techniques for video shot editing in H.264/AVC. J. Vis. Commun. Image R. 29 (2015) 2015: p. 103–113.

[33] Bestagini, P., et al. Local tampering detection in video sequences. in 2013 IEEE 15th International Workshop on Multimedia Signal Processing (MMSP). 2013. IEEE.

[34] D' Amiano, L., et al., A Patchmatch-based Dense-field Algorithm for Video Copy-Move Detection and Localization. arXiv:1703.04636, 2017., 2017.

[35] Bidokhti, A. and S. Ghaemmaghami. Detection of regional copy/move forgery in MPEG videos using optical flow. in 2015 The International Symposium on Artificial Intelligence and Signal Processing (AISP). 2015. IEEE.

[36] Nguyen, T.T., et al., Deep Learning for Deepfakes Creation and Detection. arXiv preprint arXiv:1909.11573, 2019.

[37] Sabir, E., et al., Recurrent Convolutional Strategies for Face Manipulation Detection in Videos. Interfaces (GUI), 2019. 3: p. 1.

[38] Güera, D. and E.J. Delp. Deepfake video detection using recurrent neural networks. in 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). 2018. IEEE.

International Journal of Recent Research in Mathematics Computer Science and Information Technology
Vol. 9, Issue 1, pp: (1-9), Month: April 2022 – September 2022, Available at: www.paperpublications.org

[39] Li, Y., M.-C. Chang, and S. Lyu. In ictu oculi: Exposing ai created fake videos by detecting eye blinking. in 2018 IEEE International Workshop on Information Forensics and Security (WIFS). 2018. IEEE.

[40] Li, Y. and S. Lyu, Exposing deepfake videos by detecting face warping artifacts. arXiv preprint arXiv:1811.00656, 2018. 2.

[41] Afchar, D., et al. Mesonet: a compact facial video forgery detection network. in 2018 IEEE International Workshop on Information Forensics and Security (WIFS). 2018. IEEE.

[42] Nguyen, H.H., J. Yamagishi, and I. Echizen. 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). 2019. IEEE.

[43] Nguyen, X.H., et al., Learning spatio-temporal features to detect manipulated facial videos created by the deepfake techniques. Forensic Science International: Digital Investigation, 2021. 36: p. 301108.

[44] Yang, X., Y. Li, and S. Lyu. Exposing deep fakes using inconsistent head poses. in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2019. IEEE.

[45] Matern, F., C. Riess, and M. Stamminger. Exploiting visual artifacts to expose deepfakes and face manipulations. in 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW). 2019. IEEE.

[46] Koopman, M., A.M. Rodriguez, and Z. Geradts. Detection of Deepfake Video Manipulation. in Conference: IMVIP. 2018.