A Methodological Survey on Fake Data Generation in Multimedia Forensics

Santanu Das
Department of Computer Science and Engineering
Tripura University (A Central University)
Suryamaninagar, Agartala, India
santanud803@gmail.com

Sourav Dey Roy
Department of Computer Science and Engineering
Tripura University (A Central University)
Suryamaninagar, Agartala, India
souravdeyroy49@gmail.com

Mrinal Kanti Bhowmik
Department of Computer Science and Engineering
Tripura University (A Central University)
Suryamaninagar, Agartala, India
mrinalkantibhowmik@tripurauniv.ac.in

***Abstract*— Manipulation of real-world photos/ videos through various computer-aided softwares and android applications has received significant attention. Even though most of the people across the community use these softwares for their personal entertainment, it is malicious if the people use these softwares for hiding/ concealing certain contents in images/ videos from criminal activity (i.e., forgers). Therefore, there is a huge demand for developing computer-aided systems for identifying and locating these forged regions in the images/ videos. In the present scope of this paper, we provided an extensive survey on procedures and tools for fake/ forged dataset generation. It also presents a survey on various benchmark datasets used by the researchers for fake/ forgery detection. The presented survey can be a useful contribution for the research community to develop a new method/ model for forgery detection thereby overcoming the limitations of the state-of-the-art methods.**

Keywords*— Digital Media, Forgery/ Fake, Datasets, Fake Dataset Generation Tools, Methodological Review.

I. INTRODUCTION

In the era of digitization, digital media (i.e., image, video, and voices) plays a significant role for the people to carry large amounts of information. The use of image/ video data for various purposes promotes the advancements of sophisticated editing softwares such as Photoshop [1], Flimora [2], BeautyCam[3], GIMP [4], OpenShot [5], and other android applications. With the progress of various image and video editing softwares can be considered as a two-handed sword. In one aspect, it enables the enhancement of photos/ videos by boosting the public to share their thoughts on photo editing (i.e., adjusting the brightness of the images for uploading in various social profiles; replacing the background in wedding photos). In another aspect, it is much common to forge/ fake the information of images/ videos without providing any perceptible evidence. For instance, visual information in the form of videos and images plays an important role in numerous real world applications and can be observed in news media, medical applications, education, scientific research, criminal inquiries, etc. [6]. Due to the availability of sophisticated editing softwares, forgers can efficiently spread rumors thereby creating forged images/ videos which may provide negative impacts on the society.

- There may be a situation where government officials or some well renowned celebrities have visited a place for some social/ political activities. But intentionally media/ trollers (purposely says controversial) may create some fake scenes about the person and negative impacts may arise among the people and reduce the peace and harmony of the state/ country.
- Moreover, in real world scenarios, it can be found that CCTVs are installed by governments in various public places including smart city and campus for security and surveillance. In certain situations, there is a chance for intruders to perform illegal activities. Due to this reason, when such kinds of CCTV footages are presented as evidence against certain criminal activities to the courts or forensic departments, these videos may be forged (i.e., altered) and will be very difficult to detect in naked eyes due to the high contrast loss.

Considering the above situations, a well-known real-world example is the fabricated news flash published in [7] where outlet depicts a civilian killed by U.S. Army Strykers on a major roadway. Lithuanian officials denounced the photo as an attempt to divide the NATO alliance [7]. Also, fabricated news in terms of digital photos has been published in Malaysian dailies where it has been depicted that Jeffrey Wong Su En receiving the award from Queen Elizabeth II. However, this photo was later proven fake by the original picture which was actually Ross Brawn receiving the Order of the British Empire from the Queen [8]. Thus, with the availability of manipulation tools, it is necessary to design and develop forgery detection algorithms/ models for finding manipulation in images/ videos. During the last few decades, numerous works have been published related to designing and creating various benchmark datasets so as to solve the various aspects of fake/ forgery detection problems. Conversely, in recent days, high-level image/ video manipulation has been automated by advanced computer vision technologies. The main motive of this paper is to review the procedures adopted by the researchers for forgery/ fake/ forged dataset generation.

To meet up the specific necessities, the main contribution of this paper are:

1) A detailed survey on various open source/ private tools used by the researchers for fake dataset generation are provided.
2) A detailed description on various public and private fake detection benchmark datasets is presented in this work which clarifies the way of future research.
3) It also provides a detailed review of vision based methodologies for fake dataset creation.

The whole paper is outlined as: In Section II, survey on the various tools used by the research communities for creating fake/ forged images/ videos are elaborately described. Section III illustrates the survey on vision based methodologies for fake/ forged dataset generation. In Section IV, methodological review on fake detection algorithms has been reported. And finally, Section V concludes the paper.

978-1-6654-6658-5/22/$31.00 ©2022 IEEE

2022 IEEE Region 10 Symposium (TENSYMP) | 978-1-6654-6658-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/TENSYMP54529.2022.9864441

Authorized licensed use limited to: New York University. Downloaded on September 12,2022 at 15:47:44 UTC from IEEE Xplore. Restrictions apply.

TABLE I: IMAGE/VIDEO EDITING TOOLS USED BY THE RESEARCH COMMUNITY FOR FAKE DATASET GENERATION

Editing Tools	Tools Used by the Dataset	Purpose of the Tool	Advantages	Disadvantages	Dataset Used
Adobe Photoshop CS3, CS6 [9]	SULFA [10] CoMoFoD [20] CASIA V1 [24] VTD [29]	The tool is used in these datasets for copying certain regions/ objects from same/ different images and paste it into different locations of similar/ different images.	It is a professional image editing tool. We can adjust the color balance, temperature, etc. of any image. It supports almost all image formats and is also available for all types of operating system.	It does not allow video editing, and it is not freely available. To use this tool a paid subscription is needed.	[11]-[19] [21]-[23] [26]-[28] [30]-[32]
Adobe After Effect CS5 [33]	SULFA [10]	The tool is used in these datasets for copying certain regions/ objects from images and pasting it into different locations of images.	It is prevalent because lots of filters and advanced settings are available on it to create more realistic fake/tempered videos.	It takes extensive rendering time to save a project. It is a paid software.	[11]-[19]
ZAO[34]	--	This tool can be used to choose any favorite person's face and make a video clip with the selected image.	It is very easy to use. The tool only takes an image from the users and gives the option to choose the desired face images.	It cannot correctly generate the face images of those who are not Chinese. It might be because the tool is trained on Chinese facial datasets.	--
Reface[35]	--	It allows people to impose their face on photos, videos, and GIFs. The system will create fake photos and videos.		Sometimes it generates a distorted image.	--
Face App [36]	--	It allows humans to visualize how they will look in their old age period.	It is freely available on the internet and easy to use on a smartphone.	With this tool, we can't generate any types of fake facial expressions like other deep fake generation tools.	--
Wombo [37]	--	Using this tool, anyone can make animated or .gif files from one single image of any person, and it also allows people to swap faces.	By using it, one can generate a fake video by using anyone's face images. And it also can develop the various facial expressions from the input images.	It only allows the generation of fake animated videos of single-face images and cannot generate for images containing multiple faces.	--
FakeApp [38]	UADFV [39], FFW [46], HOHA Dataset [48]	By using this tool, anyone can swap faces among two persons.	It is freely available on the internet and easy to use. Along with face swapping, the tool also performs expression swapping.	This tool may result in distorted images with uneven representation of the face contours within the images.	[40][45], [47], [49]-[55]

II. SURVEY ON FAKE DATASET GENERATION TOOLS

With the advancement of digital technologies, fake/ forged dataset generation using the data generation tools plays a significant role in various areas of computer science. The brief summary of the tools used by the research communities for fake dataset generation are summarized in TABLE I. In [11]-[19], authors have used the SULFA video dataset for forgery detection. In the SULFA dataset, all the forged videos are made by using [9] and [33] tools. Both the tools are very easy to access and give users more flexibility to edit the image/videos. For instance, they have copied a region and pasted it into the same frame to generate fake videos by using [9]. Meanwhile they have used [33] to adjust the brightness and contrast of the tempered region so as to make the videos more realistic. In [21]-[23], authors have used the [20] dataset for experimental purposes. The [20] dataset was proposed by D. Tralic et al. to apply the copy-move forgery detection algorithms. In this dataset authors generated their forged images by Adobe Photoshop. Using this tool, they have copied a small area of an image and pasted it on another place of that image. For instance, the authors of [26]-[28] have used the CASIA V1 dataset to evaluate their image splicing detection algorithm. In the CASIA dataset spliced images are created by using Adobe Photoshop CS3, CS6 [9] tool. For instance, they have used this tool to cut a small portion of an image and paste it to another portion in different/ similar image and by using this tool the whole dataset has been designed. In [30]-[32], authors had used VTD dataset to evaluate the effectiveness of the proposed method. The VTD dataset is also designed by using Adobe Photoshop CS3. Other than these tools, several Graphical User Interface (GUI) based tools were available that are used for various purposes as per the need of the users. The [34]-[37] are GUI-based fake data generation tools. These are used for the generation of counterfeit animated face images depending on the still image of anyone. By using [34][35], one can choose any favorite person's face (i.e., celebrities or any renowned person) and make a video clip/ animated file with the selected image. However [34] has some limitations because the tools are trained on Chinese people, so it can not generate face images correctly for those who are not Chinese. In contrast, [36] allows humans to visualize how they will look in their old age period. This tool is designed for entertainment purposes and is prevalent on social media platforms. By using [37], anyone can make animated files with music from one single image of any person, and it also allows people to swap faces among two-person. Nowadays, these tools have been very trending in social media and many people make fake videos by using politician's images to spread the memes.

III. SURVEY ON VISION BASED METHODOLOGIES FOR FAKE/ FORGERY CREATION IN IMAGES AND VIDEOS

In literature, several vision based algorithms were proposed to generate forged/tempered images/ videos. These algorithms were in combination known as "DeepFake" methods. This is a deep learning oriented framework(s) designed especially for generating fake images/ videos. Using the application of vision based technologies (deep fake), the researchers designed and created various forged datasets to evaluate the forgery detection models as shown in TABLE II. In [54], Shaoanlu proposed a method using the concept of generative adversarial network (GAN) to create fake images of humans. The proposed method can predict an attention mask that helps to handle occlusion, eliminate artifacts, and produce natural skin tone. The model can generate videos in different resolutions. In [55], M. Kowalski designed a method named FaceSwap. This method takes an image of a person we want to see on our own face and locate the landmark points thereby detecting the face regions. Then the 3D model is fitted on the landmark points, the vertices are considered as the

Authorized licensed use limited to: New York University. Downloaded on September 12,2022 at 15:47:44 UTC from IEEE Xplore. Restrictions apply.

TABLE II: COMPUTER VISION BASED METHODOLOGIES PROPOSED BY THE RESEARCH COMMUNITY FOR FAKE DATA GENERATION

Author/ Year	Methodology	Purpose of the Proposed Method	Dataset Created	Used by
Shaoanlu, 2018 [54]	faceswap-GAN	The method is used for generating videos of fake/ forged faces in terms of expression manipulation.	Own Dataset	[40]-[45]
M. Kowalski, 2021 [55]	FaceSwap		FaceForensics++ [56]	
P. Korshunov et al., 2018 [57]	DeepFake		Celeb-DF [58] DFDC [59]	
T. Karras et al., 2017 [60]	Generative Adversarial Network (GAN) [61]	The method is used in these datasets for generating realistic fake face images.	CELEBA-HQ [60]	[62]-[65]
C.C. Hsu et al. [66]	BigGAN [67], SN-GAN [67], SA-GAN [68],	The method is used in these datasets for generating realistic fake face images.	Own Dataset	[69]-[73]
I.P. Freelancer et al. [74]	DeepFace Lab Tool	The method is used in these datasets for generating realistic fake face swapped images.	Own Dataset	[75]
M.Y. Liu et al. [76]	FUNIT	The method generate fake object's videos of heterogeneous classes depending on the few-shot generation	Own Dataset	[77]-[78]
A. Lattas et al. [79]	AvatarMe++	These methods are used to transform the face landmark dramatically to generate fake 3D face images	Own Dataset	-
S. Ha et al. [80]	MarioNETte		Own Dataset	[81]-[83]
Y. Deng et al., 2020 [84]	DiscoFaceGAN	The method create rendered faces virtual people with disentangled	Own Dataset	-
T. Karras et al., 2019 [85]	StyleGAN	The method generate fake human face images with stochastic variation in the generated images.	Own Dataset	[86]-[88]
J. Thies et al., 2020 [89]	Face2Face	The method creates fake face dataset in terms of expression transfer.	Own Dataset	[90]-[92]
K. Olszewski et al., 2019 [93]	TBN	The method is used for generating 3D transformation of an object's images (i.e., Chair, Car, and Human) and also generate the deformation images from authentic images.	Own Dataset	[99]

texture coordinates. Once the 3D models are rendered, the image obtained from the rendered model is assorted with the image attained from the camera/ user using feathering (alpha blending) and very simple color correction. The method is proposed for swapping two faces there by manipulating the expression of the faces. Using this method, FaceForensics++ dataset [56] is designed for the research community. For instance, in [57], P. Korshunov et al. designed a GAN-based approach for face swapping. In this method, for each pair of subjects to be swapped, dual GAN models were trained two generate two video versions (i.e., low quality (LQ) and high quality (HQ) models). In the case of LQ model, a face was generated for each of the considered frame considering a frame from the input video. After that using the segmentation algorithm, facial mask was detected and the generated face was mixed with the target face video. Further for the case of HQ model, the mixing was carried out depending on the facial landmarks alignment detected between the generated face and the original face using [64] between the generated and the authentic face. Finally, histogram normalization was carried out so as to adjust the lighting conditions. Using this method, two datasets i.e., Celeb-DF, and DFDC are designed. In [60], T. Karras et al. proposed a method using GAN [61] for generating realistic fake face images. And using this method, CELEBA-HQ [60] has been designed. In [66], C.C. Hsu et al. designed a dataset for detecting the fake images. They have used three GANs from the literature to generate high-quality fake images. These GANs are BigGAN [67], SA-GAN [68], and SN-GAN [67]. In [74], I.P. Freelancer et al. proposed a face swapping method named DeepFaceLab. The method is a combined open source system consisting of three phases sequentially i.e., extraction, training, and conversion. The method achieved photorealistic face-swapping results. In [76], M.Y. Liu et al. proposed a Few-shot UNsupervised Image-to-image Translation framework based on GAN [61] for mapping an image of a source class to a target class based on learning image-to-image translation model thereby generating fake object's videos of heterogeneous classes. For instance, A. Lattas et al. [79] and S. Ha et al. [80] proposed vision based methods to create 3D fake face images. The method [79] basically uses an end-to-end reflectance inference network for reconstruction of high-quality 3D facial geometry and reflectance from a single image. Conversely, the method [80] is a few-shot face rebuilding framework named as MarioNETte which is capable of reenacting the face of unseen targets. In [84], Y. Deng et al. proposed an approach named DiscoFaceGAN which generates face images with DISentangled, precisely-COntrollable latent representations for identity of non-existing people, expression, pose, and illumination. In [85], T. Karras et al. proposed a style based GAN for generating fake images. In [89], J. Thies et al. proposed an approach for facial expression transfer/ manipulation in monocular target video sequence (i.e., videos from YouTube). In [93], K. Olszewski et al. proposed a novel network to attain 3D manipulation of image content and named it as Transformable Bottleneck Network (TBN). The proposed network utilizes one or more images to encode volumetric bottlenecks and combined in an output view coordinate frame. Further, transformed bottlenecks are then decoded to create state-of-the-art novel views, as well as reconstruct 3D geometry thereby permitting imaginative manipulations.

IV. SURVEY ON FAKE/ FORGED DATASETS

In the literature, several fake image/video datasets are proposed for developing methodologies for fake detection. The brief summary of the available dataset used for development of forgery detection algorithms/ models are summarized in TABLE III. In [95], P. Kwon et al. designed a large-scale dataset consisting of synthesized and real videos of Korean subjects. The proposed dataset contains real video and 175,776 fake video. The faking of this dataset is done in terms of face swapping, and face reenactment. In [96], V. Vinolin et al. designed a face video dataset, which contains 100 real and 100 fake face images. Each video of this dataset contains 3600 frames. In [56], A. Rossler et al. have designed a fake facial dataset for expression manipulation detection using deep fakes. The dataset contains 1000 authentic and 3000 fake face videos. In [97], H. Dang et al. designed a Diverse Fake Face Dataset (DFFD). The dataset contains 58703 real and 240336 fake images. Moreover, the dataset contains 1000 real and 3000 fake videos. In this dataset, a wide

Authorized licensed use limited to: New York University. Downloaded on September 12,2022 at 15:47:44 UTC from IEEE Xplore. Restrictions apply.

TABLE III: SURVEY ON FAKE DATASETS USED BY THE RESEARCH COMMUNITY FOR FAKE IMAGE/VIDEO DETECTION

		Dataset	Years	Statistics	Tool Category	Forged Category
Human Face Based Datasets	Videos	KoDF [95]	2021	62,166(RC) 175,776(FC)	Deep Fakes	Face Swapping, Face Reenactment
		Vinolin et al. [96]	2020	100 (AF); 100 (FF)	Traditional Tool	Object (Face) Spliced
		FaceForensics++ [56]	2019	1000 (AV); 3000 (FV)	Deep Fakes	Facial Expression Manipulation
		DFFD [97]	2019	1000 (AV); 3000 (FV)	Deep Fakes	Face Manipulation
		Celeb-DF [58]	2019	590 (AV); 5639 (FV)	Deep Fakes	Facial Expression Manipulation
		DFDC [59]	2019	4119 (FV)	Deep Fakes	Face Swapping
		UADFV [39]	2018	49 (AV); 49 (FV)	Deep Fakes	Eye Blinking Manipulation
		Deepfake Database [57]	2018	620 (FV)	Deep Fakes	Face Swapping
	Images	DSO-1 [98] DSI-1 [98]	2013	100 (AI); 100 (FI) 25 (AI);25 (FI)	Traditional Tool	Image Splicing
		DFFD [97]	2020	58703 (AI); 240336 (FI)	Deep Fakes	Face Manipulation
		iFakeFaceDB [11]	2020	87000 (FF)	Deep Fakes	Face Manipulation
		100K-Faces [99]	2019	100,000 (FF)	Deep Fakes	Face Manipulation
Object Based Datasets	Videos	VTD [29]	2016	33 (AV); 33 (FV)	Traditional Tool	Frames Splicing, Frames Swapping; and Copy Move Forgery
		KTH [100]	2016	NP	Traditional Tool	Frame Insertion and Deletion
		CVIP [41]	2015	160 (FV)	Traditional Tool	Copy Move Forgery
		SULFA [10]	2012	150 (TV)	Traditional Tool	Copy Paste or Cloning
	Images	RAISE [102]	2015	8156 (TI)	Traditional Tool	Image Splicing
		CASIA V2 [24]	2013	7200 (AI); 5,123 (FI)	Traditional Tool	Object and Region Splicing
		CASIA V1.0 [24]	2013	933 (AI); 921 (FI)	Traditional Tool	Object and Region Splicing
		CoMoFoD [20]	2013	13520 (TI)	Traditional Tool	Copy Move Forgery

AF- Authentic Face; FF- Forged Face; AV- Authentic Video; FV- Forged Video; AFR- Authentic Frame; FFR- Forged Frame; TV- Total Videos; AI- Authentic Image; Fi- Forged Image; TI- Total Image; RC- Real Clips; FC- Fake Clips

variety of face manipulation has been done including swapping of expression and identity, manipulation of attribute, and synthesizing entire human ace. In [58], Y. Li et al. designed a high-quality deep fake video dataset in terms of facial expression manipulation and given the name as celeb-DF. The dataset contains 590 real videos and 5639 fake videos. In [59], B. Dolhansky et al. designed a Deepfake Detection Challenge (DFDC) dataset incorporating the concept of deep fake method with 74% male and 26% female face samples. The dataset contains 4119 videos in terms of face swapping. In [39], Y. Li et al. designed a UADFV dataset which is a collection of deep fake generated videos and their corresponding real videos in terms of eye blinking manipulation in human faces. This dataset contains 49 real and 49 fake videos. In [57], P. Korshunov et al. created a fake face video dataset i.e., Deepfake Database using GAN guided face swapping algorithm. To design this dataset, the authors have taken video from VidTIMIT database as an original data which contains 10 videos with 43 subjects. For 16 pairs of subjects, the dataset contains 620 fake videos. In [98], T. Carvalho et al. designed a fake dataset by adding one or more persons in the original image. The dataset has two versions. The first version has 100 authentic and 100 fake images and the second version of the dataset contains 25 original and 25 fake images. In [11], J.C. Neves et al. designed an iFakeFaceDB dataset for studying the robustness of face manipulation detection algorithms. The dataset contains 87,000 synthetic face images. In [100], 100K-Faces dataset is available for study related to face manipulation. The dataset contains 100,000 forged face images. In [29], O. I.Al-Sanjary et al. have designed a Video Tampering Dataset (VTD) comprising a total of 33 authentic and 33 fake videos. In this dataset, authors have included three types of manipulation operations (i.e., splicing frames, swapping frames, and copy-move). In [100], J. Chao et al. have used TRECVID Content Based Copy Detection (CBCD) scripts to insert and delete frames in a video. In [41], E. Ardizzone et al. design a CVIP dataset which contains 160 forged videos. The dataset is designed for the study of copy move forgery detection algorithms. In [10], G. Qadir et al. generated a forged video dataset named as SULFA dataset and they have added the concept of copy-move forgery in all the considered video clips. The dataset contains total 150 videos. For instance, D.T. Dang-Nguyen et al. [102] designed a fake dataset for the forensics community. The dataset comprises of 8156 spliced images. In [24], J. Dong et al. designed CASIA tampered image dataset which contains two versions CASIA V2 and CASIA V1.0. The dataset is designed for object and region splicing. The CASIA V2 dataset contains 7200 authentic and 5123 forged images. Conversely, CASIA V1.0 contains 933 authentic and 921 forged images. In [20], D. Tralic et al. proposed a CoMoFoD dataset to detect copy-move forgery which contains 260 fake images.

V. CONCLUDING REMARKS

In this paper, we have reviewed various research works on generation of fake/ forged datasets and also provided methodological reviews on fake/ forgery generation techniques. As reviewed, it has been observed that although various datasets are designed for developing fake/ forgery detection methods but these datasets does not provide the variance of small and large objects forged/ faked in the images/ frames. Therefore, there is a need for designing a large annotated object based forged datasets in both indoor and outdoor environments.

REFERENCES

[1] R. Concepcion, “Adobe Photoshop and Lightroom Classic CC Classroom in a Book (2019 release),”Adobe Press, 2019.

[2] The Best Video Editor Alternative to Windows Movie Maker & iMovie. [Online]. Available: https://filmora.wondershare.net/

[3] W. Zhang, “Smartphone photography in urban China,” International Journal of Humanities and Social Sciences, Vol. 11, No. 1, pp. 231-239, 2017.

[4] GNU Image Manipulation Program (GIMP). [Online]. Available: https://www.gimp.org/

[5] Openshot Video Editor - Easy-To-Use Alternative. [Online]. Available: https://www.movavi.com/

[6] Y. Liu, X. Zhu, X. Zhao, and Y. Cao, “Adversarial learning for constrained image splicing detection and localization based on atrous convolution,” IEEE Transactions on Information Forensics and Security, Vol. 14, No. 10, pp. 2551-2566, 2019.

[7] Ever heard of ‘deep fake’ technology? The phony audio and video tech could be used to blackmail US troops. Available [online]. https://www.militarytimes.com/news/your-air-force/2018/07/19/ever-heard-of-deep-fake-technology-the-phony-audio-and-video-tech-could-be-used-to-blackmail-us-troops/

Authorized licensed use limited to: New York University. Downloaded on September 12,2022 at 15:47:44 UTC from IEEE Xplore. Restrictions apply.

[8] J.A. Redi, W. Taktak, and J.L. Dugelay, "Digital image forensics: a booklet for beginners," Multimedia Tools and Applications, Springer, Vol. 51, No. 1, pp. 133-162, 2011.

[9] Adobe Photoshop CS3 Update, [Online]. Available: https://adobe-photoshop-cs3-update.en.softonic.com/.

[10] G. Qadir, S. Yahaya, and A. T. Ho, "Surrey university library for forensic analysis (SULFA) of video content," pp. 121-121, 2012.

[11] J. C. Neves, R. Tolosana, R. Vera-Rodriguez, V. Lopes, H. Proença, and J. Fierrez, "Ganprintr: Improved fakes and evaluation of the state of the art in face manipulation detection", IEEE Journal of Selected Topics in Signal Processing, Vol. 14, No. 5, pp. 1038-1048, 2020.

[12] S. Fadl, A. Megahed, Q. Han, and L. Qiong, "Frame duplication and shuffling forgery detection technique in surveillance videos based on temporal average and gray level co-occurrence matrix," Multimedia Tools & Applications, Springer, Vol. 79, 2020.

[13] M. Aloraini, M. Sharifzadeh, C. Agarwal, and D. Schonfeld, "Statistical sequential analysis for object-based video forgery detection," Electronic Imaging, Spie, No. 5, pp. 543-1, 2019.

[14] J. L. Zhong, C. M. Pun, and Y. F. Gan, "Dense moment feature index and best match algorithms for video copy-move forgery detection," Information Sciences, Elsevier, Vol. 537, pp. 184-202, 2020.

[15] G. Chittapur, S. Murali, and B. S. Anami, "Copy Create Video Forgery Detection Techniques Using Frame Correlation Difference by Referring SVM Classifier," International Journal of Computer Engineering In Research Trends (IJCERT), 2019.

[16] J. Bakas, R. Naskar, and R. Dixit, "Detection and localization of inter-frame video forgeries based on inconsistency in correlation distribution between Haralick coded frames," Multimedia Tools and Applications, Elsevier, Vol. 78, No. 4, pp. 4905-4935, 2019.

[17] Q. Li, R. Wang, and D. Xu, "An inter-frame forgery detection algorithm for surveillance video," Information, Vol. 9, No. 12, pp. 301, 2018

[18] S. M. Fadl, Q. Han, and Q. Li, "Authentication of surveillance videos: detecting frame duplication based on residual frame," Journal of forensic sciences, Wiley-Blackwell, Vol. 63, No. 4, pp. 1099-1109, 2018.

[19] Y. Liu, and T. Huang, "Exposing video inter-frame forgery by Zernike opponent chromaticity moments and coarseness analysis," Multimedia Systems, Vol. 23, No. 2, pp. 223-238, 2017

[20] D. Tralic, I. Zupancic, S. Grgic, and M. Grgic, "CoMoFoD—New database for copy-move forgery detection," In Proceedings ELMAR-2013, IEEE, pp. 49-54, 2013.

[21] Y. Liu, Q. Guan, and X. Zhao, "Copy-move forgery detection based on convolutional kernel network," Multimedia Tools and Applications, Springer, Vol. 77, No. 14, pp. 18269-18293, 2018.

[22] Y. Liu, C. Xia, X. Zhu, and S. Xu, "Two-Stage Copy-Move Forgery Detection with Self Deep Matching and Proposal SuperGlue," arXiv preprint arXiv:2012.08697, 2020.

[23] Y. Zhu, C. Chen, G. Yan, Y. Guo, and Y. Dong, "AR-Net: Adaptive attention and residual refinement network for copy-move forgery detection," IEEE Transactions on Industrial Informatics, Vol. 16, No. 10, pp. 6714-6723, 2020.

[24] J. Dong, W. Wang and T. Tan, "Casia Image Tampering Detection Evaluation Database," in Proc. 2013 IEEE China Summit and International Conference on Signal and Information Processing, IEEE, pp. 422-426, 2013.

[25] R.Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, "Deepfakes and beyond: A survey of face manipulation and fake detection," Information Fusion, Vol. 64, pp. 131-148, 2020.

[26] A. El-Latif, I. Eman, A. Taha, and H. H. Zayed, A Passive Approach for Detecting Image Splicing using Deep Learning and Haar Wavelet Transform. International Journal of Computer Network & Information Security, 11(5).

[27] R. Salloum, Y. Ren, and C. C. J. Kuo, "Image splicing localization using a multi-task fully convolutional network (MFCN)," Journal of Visual Communication and Image Representation, Elsevier, Vol. 51, pp. 201-209, 2018.

[28] Y. Rao, and J. Ni, "A deep learning approach to detection of splicing and copy-move forgeries in images" In IEEE International Workshop on Information Forensics and Security (WIFS), IEEE, pp. 1-6, 2016.

[29] O. I. Al-Sanjary, A. A. Ahmed, and G. Sulong, "Development of a video tampering dataset for forensic investigation," Forensic science international, Elsevier, Vol. 266, pp. 565-572, 2016.

[30] H. D. Panchal, and H. B. Shah, "Video tampering dataset development in temporal domain for video forgery authentication," Multimedia Tools and Applications, Springer, Vol. 79, No. 33, 2020.

[31] Q. Bai, Z. Wu, S. Sclaroff, M. Betke, and Monnier, C., "Randomized ensemble tracking," In Proceedings of the IEEE International Conference on Computer Vision, pp. 2040-2047, 2013.

[32] J. Kwon, and K. M. Lee, "Visual tracking decomposition", In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, pp. 1269-1276, 2010.

[33] Adobe After Effects CS5 Free Download. [Online]. https://webintopc.com/adobe-after-effects-cs5-free-download/.

[34] 'ZAO App', [online]. Avaiable: https://zaodownload.com/

[35] 'RefaceApp', [online]. Available: https://reface.app/

[36] 'FaceApp', [online]. Available: https://www.faceapp.com/

[37] 'WOMBO', [online]. Available: https://www.wombo.ai/

[38] 'FakeApp', [online]. Available: https://www.malavida.com/en/soft/fakeapp/.

[39] X., Yang, Y. Li, and S. Lyu, (2019, May). Exposing deep fakes using inconsistent head poses. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 8261-8265). IEEE.

[40] F. Matern, C. Riess, and M. Stamminger, "Exploiting visual artifacts to expose deepfakes and face manipulations" In 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW), IEEE, pp. 83-92, 2019.

[41] E. Ardizzone, and G. Mazzola, "A tool to support the creation of datasets of tampered videos," in Proc. International Conference on Image Analysis and Processing, Springer, Cham, pp. 665-675, 2015.

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

[43] P. Zhou, X. Han, V. I. Morariu, and L. S. Davis, "Two-stream neural networks for tampered face detection," In IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, pp. 1831-1839, 2017.

[44] H. H. Nguyen, F. Fang, J. Yamagishi, and I. Echizen, "Multi-task learning for detecting and segmenting manipulated facial images and videos," arXiv preprint arXiv:1906.06876, 2019.

[45] R. Tolosana, S. Romero-Tapiador, J. Fierrez, and R. Vera-Rodriguez, "Deepfakes evolution: Analysis of facial regions and fake detection performance," In International Conference on Pattern Recognition, Springer, pp. 442-456, 2021.

[46] A. Khodabakhsh, R. Ramachandra, K. Raja, P. Wasnik, and C. Busch, "Fake face detection methods: Can they be generalized?" In 2018 international conference of the biometrics special interest group (BIOSIG), IEEE, pp. 1-6, 2018.

[47] A. Khodabakhsh, R. Ramachandra, K. Raja, P. Wasnik, and C. Busch, "Fake face detection methods: Can they be generalized?", In 2018 international conference of the biometrics special interest group (BIOSIG), IEEE, pp. 1-6, 2018.

[48] I. Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld, "Learning realistic human actions from movies," In 2008 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 1-8, 2008.

[49] D. Güera, 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), IEEE, pp. 1-6, 2018.

[50] I. Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld, "Learning realistic human actions from movies," In 2008 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 1-8, 2008.

[51] K. Rapantzikos, Y. Avrithis, and S. Kollias,"Dense saliency-based spatiotemporal feature points for action recognition," In 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, pp. 1454-146, 2009.

[52] M. Raptis, I. Kokkinos, and S. Soatto, "Discovering discriminative action parts from mid-level video representations," In 2012 IEEE conference on computer vision and pattern recognition, IEEE, pp. 1242-1249, 2012.

[53] M. Raptis, and S. Soatto, "Tracklet descriptors for action modeling and video analysis," In European conference on computer vision, Springer, pp. 577-590, 2010.

[54] Shaoanlu, "Shaoanlu/Faceswap-Gan: A denoising autoencoder + adversarial losses and attention mechanisms for face swapping," GitHub. [Online]. Available: https://github.com/shaoanlu/faceswap-GAN. [Accessed: 28-Aug-2021]

[55] M. Kowalski, "MarekKowalski/FaceSwap: 3D face Swapping implemented in Python," GitHub. [Online]. Available: https://github.com/MarekKowalski/FaceSwap.

[56] A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "Faceforensics++: Learning to detect manipulated facial images," In Proceedings of the IEEE/CVF International Conference on Computer Vision, IEEE pp. 1-11, 2019.

[57] P. Korshunov, and S. Marcel, "Deepfakes: a new threat to face recognition? Assessment and detection," arXiv preprint arXiv:1812.08685, 2018.

[58] Y. Li, X. Yang, P. Sun, H. Qi, and S. Lyu, "Celeb-df: A large-scale challenging dataset for deepfake forensics," In Proceedings of the

Authorized licensed use limited to: New York University. Downloaded on September 12,2022 at 15:47:44 UTC from IEEE Xplore. Restrictions apply.

IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3207-3216, 2020.

[59] B. Dolhansky, R. Howes, B. Pflaum, N. Baram, and C. C. Ferrer, (2019). The deepfake detection challenge (dfdc) preview dataset. arXiv preprint arXiv:1910.08854.

[60] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," arXiv preprint arXiv:1710.10196.

[61] I. Goodfellow, J. Pouget-Abadie, M. , Mirza, B. Xu, D. Warde-Farley, S. Ozair, and Y. Bengio, Generative adversarial nets. Advances in neural information processing systems, 27, 2014.

[62] H. Mo, B. Chen, and W. Luo, Fake faces identification via convolutional neural network. In Proceedings of the 6th ACM workshop on information hiding and multimedia security (pp. 43-47), June, 2018.

[63] Y. Choi, Y. Uh, J. Yoo, and J. W. Ha, Stargan v2: Diverse image synthesis for multiple domains. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 8188-8197), 2020.

[64] P. Esser, R. Robin, and O. Björn, "A note on data biases in generative models." arXiv preprint arXiv:2012.02516 (2020).

[65] P. Patel, N. Kumari, M. Singh, and B. Krishnamurthy, "Lt-gan: Self-supervised gan with latent transformation detection," In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 3189-3198), 2021.

[66] C. C. Hsu, Y. X. Zhuang, and C. Y. Lee, "Deep fake image detection based on pairwise learning. Applied Sciences", 10(1), 370, 2020.

[67] A. Brock, J. Donahue, and K. Simonyan, "Large scale GAN training for high fidelity natural image synthesis," arXiv preprint arXiv:1809.11096, 2018.

[68] H. Emami, M. Dong, S. P. Nejad-Davarani, and C. K. Glide-Hurst, "SA-GAN: Structure-Aware GAN for Organ-Preserving Synthetic CT Generation," In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 471-481). Springer, Cham, September, 2021.

[69] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, and L. Fei-Fei, "Imagenet large scale visual recognition challenge. International journal of computer vision, 115(3), 211-252, 2015.

[70] C. Tai, T. Xiao, Y. Zhang, and X. Wang, "Convolutional neural networks with low-rank regularization", arXiv preprint arXiv:1511.06067, 2015.

[71] S. Oh, M. Kim, D. Kim, M. Jeong, and M. Lee, "Investigation on performance and energy efficiency of CNN-based object detection on embedded device," In 4th International Conference on Computer Applications and Information Processing Technology (CAIPT) (pp. 1-4). IEEE, August 2017.

[72] S. Pattanayak, S. Nag, & S. Mittal, "CURATING: A multi-objective based pruning technique for CNNs," Journal of Systems Architecture, 116, 102031, 2021.

[73] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, "Spectral normalization for generative adversarial networks," arXiv preprint arXiv:1802.05957, 2018.

[74] I. Perov, G. Daiheng, C. Nikolay, L. Kunlin, M. Sugasa, U. Chris, et al. "DeepFaceLab: Integrated, flexible and extensible face-swapping framework." arXiv preprint arXiv:2005.05535, 2020.

[75] T. J. De Carvalho, C. Riess, E. Angelopoulou, H. Pedrini, and A. de Rezende Rocha, "Exposing digital image forgeries by illumination color classification," IEEE Transactions on Information Forensics and Security, 8(7), 1182-1194, 2013.

[76] M. Y. Liu, X. Huang, A. Mallya, T. Karras, T. Aila, J. Lehtinen, and J. Kautz, "Few-shot unsupervised image-to-image translation," In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 10551-10560), 2019.

[77] L. Barras, S. Chassot, and D. F. N. Silva, "Few-Shot Unsupervised Image-to-Image Translation on complex scenes," arXiv preprint arXiv:2106.03770, 2021.

[78] K. Saito, K. Saenko, and M. Y. Liu, "Coco-funit: Few-shot unsupervised image translation with a content conditioned style encoder," In European Conference on Computer Vision (pp. 382-398). Springer, Cham, August, 2020.

[79] A. Lattas, S. Moschoglou, S. Ploumpis, B. Gecer, A. Ghosh, and S. P. Zafeiriou, "AvatarMe++: Facial Shape and BRDF Inference with Photorealistic Rendering-Aware GANs," IEEE Transactions on Pattern Analysis & Machine Intelligence, (01), 1-1, 2021.

[80] S. Ha, M. Kersner, B. Kim, S. Seo, and D. Kim, "Marionette: Few-shot face reenactment preserving identity of unseen targets," In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 34, No. 07, pp. 10893-10900), April, 2020.

[81] G. Ariyanto, and M. S. Nixon, "Marionette mass-spring model for 3d gait biometrics," In 2012 5th IAPR International Conference on Biometrics (ICB) (pp. 354-359). IEEE, March, 2012.

[82] X. Wu, Z. Feng, W. Fan, J. Gao, and Y. Yu, "Detecting marionette microblog users for improved information credibility," In Joint European Conference on Machine Learning and Knowledge Discovery in Databases (pp. 483-498). Springer, September, 2013.

[83] M. Ajallooeian, A. Borji, B. N. Araabi, M. N. Ahmadabadi and H. Moradi, "Fast hand gesture recognition based on saliency maps: An application to interactive robotic marionette playing," In RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication (pp. 841-847). IEEE, October, 2009.

[84] Y. Deng, J. Yang, D. Chen, F. Wen, and X. Tong, "Disentangled and controllable face image generation via 3d imitative-contrastive learning," In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5154-5163), 2020.

[85] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 4401-4410), 2019.

[86] A., Tewari, M. Elgharib, G. Bharaj, F. Bernard, H. P. Seidel, P. Pérez, and C. Theobalt, "Stylerig: Rigging stylegan for 3d control over portrait images," In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6142-6151), 2020.

[87] L. Li, J. Bao, H. Yang, D. Chen, and F. Wen, "Faceshifter: Towards high fidelity and occlusion aware face swapping," arXiv preprint arXiv:1912.13457, 2019.

[88] Y. Nirkin, Y. Keller, and T. Hassner, "Fsgan: Subject agnostic face swapping and reenactment," In Proceedings of the IEEE/CVF international conference on computer vision (pp. 7184-7193), 2019.

[89] J. Thies, M. Zollhofer, M. Stamminger, C. Theobalt, and M. Nießner, "Face2face: Real-time face capture and reenactment of rgb videos," In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2387-2395), 2016.

[90] D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, "Mesonet: a compact facial video forgery detection network," In 2018 IEEE international workshop on information forensics and security (WIFS) (pp. 1-7). IEEE, December, 2018.

[91] S. Tariq, S. Lee, and S. S. Woo, "A convolutional LSTM based residual network for deepfake video detection," arXiv preprint arXiv:2009.07480, 2020.

[92] P. Kumar, M. Vatsa, and R. Singh, "Detecting face2face facial reenactment in videos," In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 2589-2597), 2020.

[93] K. Olszewski, S. Tulyakov, O. Woodford, H. Li, and L. Luo, "Transformable bottleneck networks," In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 7648-7657), 2019.

[94] J. Ren, M. Chai, O. J. Woodford, K. Olszewski, and S. Tulyakov, "Flow guided transformable bottleneck networks for motion retargeting," In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10795-10805), 2021.

[95] P. Kwon, J. You, G. Nam, S. Park, and G. Chae, "Kodf: A large-scale korean deepfake detection dataset," In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 10744-10753), 2021.

[96] V. Vinolin, and M. Sucharitha, "Dual adaptive deep convolutional neural network for video forgery detection in 3D lighting environment," The Visual Computer (2020): 1-22, Springer, 2021.

[97] H. Dang, F. Liu, J. Stehouwer, X., Liu, and A. K. Jain, "On the detection of digital face manipulation," In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern recognition (pp. 5781-5790), 2020.

[98] T. Carvalho, F. A. Faria, H. Pedrini, R. D. S. Torres, and A. Rocha, "Illuminant-based transformed spaces for image forensics," IEEE transactions on information forensics and security, 11(4), 720-733, 2015.

[99] 100,000 Faces Generated by AI, 2018. [Online]. Available: https://generated.photos/

[100] J. Chao, 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 (pp. 267-281), 2012.

[101] L. Su, T. Huang, and J. Yang, "A video forgery detection algorithm based on compressive sensing," Multimedia Tools and Applications, 74(17), 6641-6656, 2015.

[102] D. T. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, "Raise: A raw images dataset for digital image forensics," In Proc. 6th ACM multimedia systems conference, ACM, pp. 219-224, 2015.

Authorized licensed use limited to: New York University. Downloaded on September 12,2022 at 15:47:44 UTC from IEEE Xplore. Restrictions apply.