

Sixth International Conference on Futuristic Trends in Networks and Computing Technologies (FTNCT06) held in
Uttarakhand, India

On Machine Learning and Deep Learning based Deepfake Generation and Detection

Mohd Tahir Irfan^a, Bhavna Arora^b, Neha Sandotra^c, Abrar Ahmed Raza^d

^bDepartment of Computer Science Engineering, Central University of Jammu-181143, J&K, India

^{a,c,d}Department of Computer Science & IT, Central University of Jammu-18114, J&K, India

Abstract

With the advancement of artificial intelligence, deepfakes have evolved into a potent tool that allows the developer to manipulate images or audios that can lead to defamation or any other kind of security threat. It is a cutting-edge technology that uses deep learning and machine learning techniques which gives the user enormous power to create deep fake media for both entertainment and malicious purposes that may result in high impact in real life scenarios. Hence, recently the research communities have been increasingly interested in the development of approaches for detecting deepfakes as the trust on the media available online comes under dilemma. In this paper, a comprehensive overview of deepfake technology with its pros and cons, followed by deepfake generation methods like Encoder-Decoder and GAN is discussed. The benchmark datasets with the open-source tools for deepfake generation have also been discussed in detail. How the face manipulation techniques like Face-Swap, Face-Synthesis, Face-Attribute-Manipulation and Face-Re-enactment are used is also a part of this study. Additionally, it offers a comparison of past research on the identification of deepfake images and videos which are applying deep-learning and machine-learning algorithms. The research gaps of this technology and how this can be implemented for further research, perspective, and insights of the same have also been given. An evaluation on the machine-learning and deep-learning based detection models for fake images, videos, audios, and multimodal content has also been explored and presented in this paper.

© 2025 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the Sixth International Conference on Futuristic Trends in Networks and Computing Technologies (FTNCT06)

Keywords: Deepfake Generation, Deepfake Detection, Machine Learning, Deep Learning

1. Introduction

The rapid advancement of technology and the accessibility of reasonably priced intelligent devices such as laptops, smartphones, tablets, and various types of digital cameras have led to an exponential rise in the amount of multimedia content (e.g video, images & audio). The user can share the captured content across the internet quickly and easily, due to the advancement in digital multimedia. The sharing also leads to the widespread distribution of fake news,

which may threaten public's trust (Arora 2016; Hangloo and Arora 2021, 2022). Through the usage of computer vision and deep learning technology, the latest developing technology has been revealed which allows anybody to generate and distribute incredibly realistic but fake videos, images, and even voiceovers. The spread of these fake images and videos raises many questions, making Deepfake very famous in today's technology. Deepfake is a term which is derived from the words 'Deep Learning (DL)' and 'Fake,' and it refers to a specialized photo-realistic image or video content generated by applying DL's technology. Deepfake is a technique that utilizes AI to replace a person's face in a picture or video with the face of another person in order to fool the target person into doing or saying anything. (Lee and Kim 2021a). The term deepfake was initially invented by the Reddit anonymous user account name "Deepfake" who then asserted to have developed a machine learning method to convert celebrity images into pornographic content (Zhang 2022). In the year 2018, a fake video of former president of United States Barack Obama was created and circulated in which certain words are used which he never said. Using deepfake technology, the creator of this video transmitted his facial movements to Obama's facial characteristics (Appel and Prietzel 2022). Furthermore, this technique has earlier been used during the United States 2020 election in which Joe Biden with tongue poking out can be seen. However, although the research on the ethical implications of deepfake technology highlights its potentially far-reaching consequences in politics, the most prevalent present application of deepfake technology is for pornographic purposes. Based on yearly reports Deepfake pornography is a global phenomenon with a huge following on multiple specialized websites, with women being specifically targeted. Deepfake videos are fast spreading online, with the number almost doubling and tripling in the previous 10 months. The rising commercialization of tools and services, which reduces the barrier for non-experts to build deepfakes, is fueling this surge. This demonstrated that this technology is continually growing and has the potential to mislead a large portion of the people.

1.1. Contribution and Scope of our Survey

This study primarily focuses on the technical aspects and various applications of deepfakes. Technical aspects involve the generation and detection method of deepfake. This research conducted a systematic review of the existing literature on deepfake images and videos. The following summarizes the contribution of our survey:

- Categorize the deepfake Pros and Cons and further expanded to give detailed descriptions in different scenarios.
- Discuss the benchmark dataset generated from the year 2018 to 2023 in tabular form.
- Outline the various open-source tools and AI based algorithms for deepfake generation. Also provide face manipulations techniques to generate fake images or videos.
- Discuss detailed descriptions of detection techniques of deepfake and highlights the key findings of each technique.
- Identify limitations of existing detection methods and present the future trends for this technology in coming years.

The structure of paper is organized as follows: section 2 discussed literature review. Section 3 discusses the generating techniques, whereas Section 4 discusses the detection techniques. Section 5 discussed the evaluation of detection techniques. In section 6 various research gaps and future directions have been discussed and finally our work is concluded in section 7.

2. Literature Review

Several research papers have investigated thoroughly deepfakes and categorizing methods at multiple levels (Neha and Arora 2023; Sandotra and Arora 2023). In (Tolosana et al. 2020) deepfake types are divided into four categories, which are restricted to images and videos, without the detection mechanism. In this section, we will discuss the collection of articles, journals, and various research papers. Fig. 1 graphically shows the studied articles related to the term deepfake, that are collected from statistics of Google Scholar from the year 2018 to 2023. The graph indicates that there has been more research in recent years. Although the anticipated number of deepfake publications may be lower than the real number, there is no doubt that this is an emerging trend that is developing.

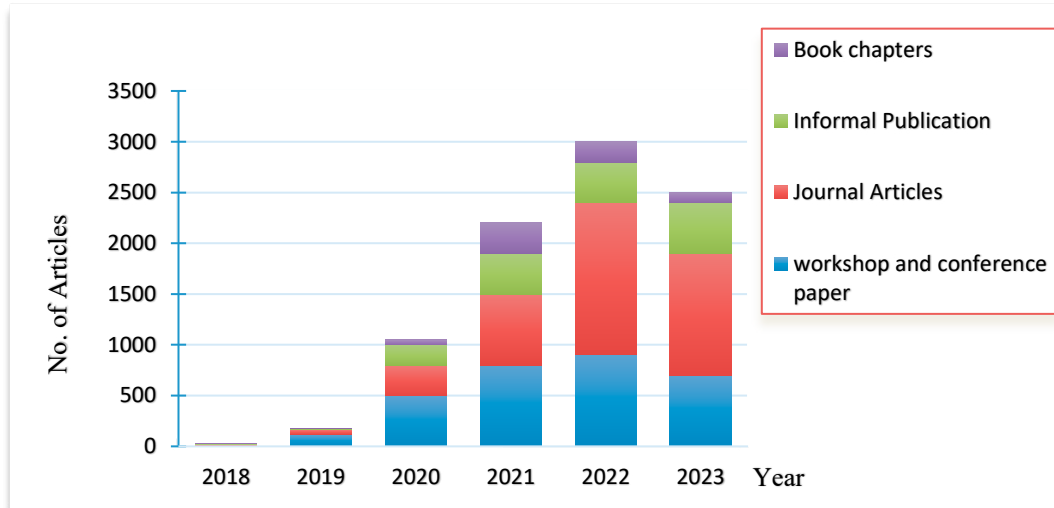


Fig 1. The number of studies on deepfakes published in Google Scholar between 2018 and 2023

2.1. Deepfake Datasets

Deep learning techniques are used in both deepfake generation and detection, and huge dataset required in training and testing process to identify deepfake efficiently. However, collecting a large deepfake dataset is difficult, especially in the case of databases containing human faces, because of privacy concerns. Fortunately, there are datasets that are publicly available. This section will cover the openly available dataset based on image and video, represented in table 1.

Table 1. Overview of Deepfake Dataset

Dataset	Released Year	Image	Video	Summary
Fake Face in the wild (FFW)(Brömme et al. 2018)	2018	✓	✗	FFW includes wide variety of fake images that are generated through computer graphic images, GAN, manual and automatic image-tempering methods.
Flicker-Faces-HQ (FFHQ)(Karras, Laine, and Aila 2019)	2019	✓	✗	GAN generated 70,000 face images with excellent grade resolution.
Face forensics (FF) (Rössler et al. 2018)	2019	✓	✓	Faces from 1004 original videos are represented in this dataset. The manipulation of faces is generated with the state of art face editing approach
Face forensics++ (FF++) (Rössler et al. 2019)	2019	✓	✓	1.8M images and 1000 original videos manipulated with four manipulated techniques (deepfakes, face2face, face-swap, and neural textures).
Celeb-DF(Li et al. n.d.)	2019	✓	✓	The advanced deepfake synthesis algorithm is used to generate deepfake videos. It has 590 authentic videos and 5639 deepfake videos.
140k real and fake faces(Anon n.d.-b)	2019	✓	✗	Contains 70k real faces from flicker dataset and 70k face are sampled from 1M fake faces that are generated by Style-GAN.
Deepfake MNIST+ (Huang et al. n.d.)	2021	✓	✓	Generated by state-of-art image animation generator. The first large-scale face animation video dataset that contains 10k fake and 10k realistic animation videos.
Gender based deepfake dataset (Nadimpalli and Rattani 2022)	2022	✓	✓	Deep forensics -1.0, FF++, Celeb-DF were used to create a gender-based deepfake dataset (GBDF). Comprised 10,000 real and fake videos.

3. Deepfake Generation Techniques

Deepfake creation is a unique media tempering technology which tackles the key limitations of previous forgery generation systems by removing manipulation traces or fingerprints. Deepfakes have grown in popularity due to the great quality of the manipulated videos and the ease of usage for a wide range of users, from beginners to experts. These apps are generally built using deep learning techniques. To generate deepfake two methods are involved (i) the Encoder-Decoder method and (ii) the Generative adversarial networks (GAN) method.

3.1. Encoder-Decoder Method

In this method two encoder-decoder pairs are required to swap faces between the source images and target images. The encoder changes the input, into a hidden latent representation, from which decoder attempts to reconstruct the data back to its original form. Each pair of encoder-decoders is trained on a distinct set of images, while the parameters of the encoder are shared by two network pairs. Or to put it another way, both pairs use the same encoder network. This technique enables the shared encoder to recognize the similarities among two groups of the images of face, which is comparatively straightforward given that human faces typically have comparable features such as nose shape, lips, eyes, and mouth configuration. (Masood et al. 2022a).

3.2. Generative adversarial network (GAN) Method

The Architecture of GAN consists of two neural networks, one of which is a generator and the another one is a discriminator. The fake images are created using a generator, and the discriminator verifies the legitimacy of the images (Real or fake). In other words, the generative model is like a group of counterfeiters attempting to make counterfeit currency and used it without detection, whereas the discriminative model is similar to police attempting to identify that counterfeit currency. The competition in this game drives both sides to develop their techniques until the counterfeits or forgeries are undetectable from the actual products (Paul n.d.). The two components physically play the roles of a forger and a detective, and they are operationally competitive. This method is applied to various works such as Faceapp, FaceswapGAN, FSGAN and Wave2lip.

Given a set of real images x , which has a P_{data} distribution, the main objective of the generator G is to generate an image $G(z)$ that resembles real images x , where z is a noise signal and P_z is the distribution. The primary goal of the discriminator D is to exactly identify the true image x and the image which is generated by G . The discriminator D is trained so that it can increase its classification accuracy, i.e so that it can maximize $D(x)$ which tells us the probability that x is a real image instead of a fake one generated by G . Similarly, G is trained such that it minimize the probability that its output will be classified as a synthetic image by D , i.e to minimize $1-D(G(z))$. The generative and the discriminative model both plays a min-max game where G tries to maximize the probability of D and wanted D to making a mistake, and mathematically equation written as

$$\min \max V(G, D) \quad (1)$$

The value function of min-max game $V(G, D)$ is represented as

$$V(G, D) = E_{x \sim P_{data}(x)} [\log D(x)] + E_{(z \sim P_z(z))} [\log (1 - D(G(z)))] \quad (2)$$

3.3. Face Manipulation Techniques

Face Manipulation are different types in images and videos. Face manipulation types are categorized into following types: Face-swap, Face-Synthesis, Face-Attribute Manipulation and Face-Reenactments.

- i. *Face-swap*: This approach is based on an auto-encoder and decoder method. In this method, an An encoder extracts the face's latent features from a picture and a decoder then reconstructs the face The first Face-Swap used FaceForensics++ (Rössler et al. 2019) dataset for the generation of deepfake.
- ii. *Face-Synthesis*: Face-Synthesis is the process of creating photorealistic representations of a human faces that may or may not be real.
- iii. *Face-Attribute Manipulation*: Face-Attribute Manipulation is adjusting an attribute-specific area while leaving irrelevant parts alone in order to change the facial features of an existing sample.
- iv. *Face-Reenactment*: This technique alters a person's facial expression by projecting the source actor's facial gestures, like eye movements and head movements onto a resultant image or video. The face manipulated deepfakes which are discussed above are depict in Fig 2.

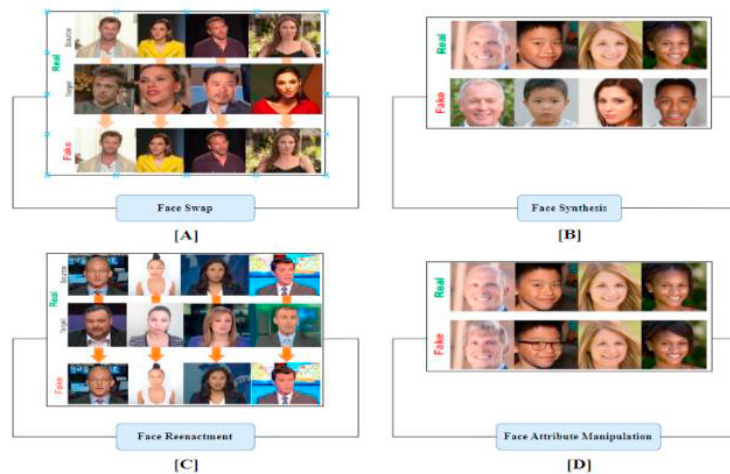


Fig 2. Samples of Deepfake images manipulated by different face manipulation techniques(Akhtar, Dasgupta, and Banerjee 2019)

4. Deepfake Detection Techniques

This section is categorized in three subsections, first section will discuss deepfake detection based on machine learning (ML), second section discuss deep learning-based detection and last section will evaluate the deepfake detection of images, videos, audios or multimodal on both ML and DL techniques. Deepfakes are dangerous to privacy, social security, and democracy (Robert Chesney and Danielle Keats Citron. n.d.). Deepfake detection approaches have been developed since the threat was firstly revealed. Early attempts for detection are reliant on handcrafted features that arising from fault and mistakes in the fake image generation process (Masood et al. 2022b). Later, ML methods were used to detect deepfakes and gives a very prominent result. DL approaches have recently been widely used in detecting deepfakes. To detect deepfake, this approach automatically extracts significant and discriminative features. Recent methods applied the DL methods to extract the discriminative features for detecting deepfakes. The Fig.3 illustrates that the DL detection approach surpasses the ML method and other methods. Table.2 presenting a summary of deepfake detection model based on ML and DL algorithms.

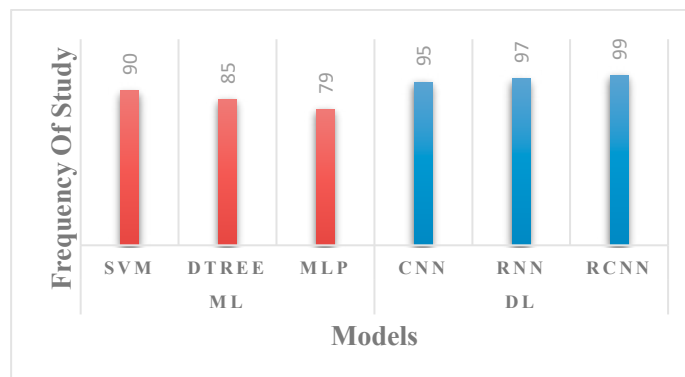


Fig 3. Illustrates that the DL detection approach surpasses the ML method and other methods

4.1. Machine Learning based Detection

Conventional ML techniques are useful for understanding the reason behind any result which can be described in human words. As ML methods solves numerous real-time problems. For deepfake detection domain ML techniques are appropriate as they have better comprehension of the data and procedures. Furthermore, changing hyper-parameters and altering model architectures are significantly easier. As the GANs models are autonomously train generative models by treating an unsupervised problem as supervised and generating photorealistic synthetic faces in

images or videos. Some ML based algorithms are designed to highlight specific irregularities which are seen in GAN-generated fake videos or images. The deepfake's primary technique is to alter or to significantly change the human faces in images or videos in order to confuse its viewers. There are several techniques to performing this type of deepfake. In most techniques trick users are modifying certain region of the face, such as the color of the eyes, rings in ears, and so on. These methods rely on a specific portion (also known as a feature) to identify or detect the altered area. To address these issues, the authors of (F. Matern n.d.) suggested a deepfake approach that combines a set of such traits. In terms of performance, ML based methods have been found to achieve a good percentage in accuracy for the detecting Deepfakes. Moreover, detection performance is fully dependent by the type of dataset, the selected features, and the orientation of the training and testing sets. The unrelated dataset affects the performance by about half, which is an unreasonable suggestion. The ML techniques have specific ML models for deepfake detection. Using a range of advanced feature selection algorithms, this approach creates a feature vector by selecting the essential features.(Uçar 2020). It then trains a classifier with this vector as input to determine whether the modified images are deep fakes or not. The ML based models used for deepfake detection are Support Vector Machine (SVM), k-Means Multilayer-Perceptron, Random-Forest, Decision-Tree, Discriminant-Analysis, Logistic-Regression and Naive-Bayes.

4.2. Deep Learning based Detection

There are numerous works on Deepfake detection in images that use deep learning-based approaches to detect certain artifacts caused by their creation pipeline. Deep learning methods, such as Convolution neural network (CNN), Recurrent neural network (RNN), Recurrent convolution neural network (RCNN), Multi-cascaded convolution neural network (MTCNN), which can automatically extract deepfake features and are broadly used in the detection of deepfake when compared to traditional machine learning models(AI-Dhabi and Zhang 2021). DL based detection approaches have the potential to improve detection accuracy. In study (Anon n.d.-a) proposed a GAN simulator that replicates aggregate GAN-image artifacts and sends them to a Deepfake classifier as input. Rahmouni et al. (Rahmouni N n.d.) present a DL technique that employs feature extraction and a CNN framework to distinguish computer-generated faked images and real images. CNN model (e.g., XceptionNet, GoogleNet, ResNet, Efficient-Net, HRNet, MobileNet, InceptionV3, DenseNet, SuppressNet,), RNN model include Bidirectional RNN, Long-term Recurrent Convolutional Neural Network model, Hierarchical Memory Network model, and Multi-task Cascaded CNNs model. DL detection methods outperform ML methods in terms of performance. The DL technique was proven to be 99 percent accurate in detecting deepfakes, However, due to the unrelated dataset, this technique has still yet to set a detection benchmark.

Table 2. Deepfake detection using Machine Learning (ML) and Deep Learning (DL) techniques

Year	Author	Dataset	Techniques		Key Findings	Accuracy
			ML	DL		
2021	Hasin Shahed Shad et al.(Shad et al. 2021)	Flicker Dataset	✗	✓	CNN is extremely successful in detecting and classifying GAN-generated images. More efficient models are required to detect deepfakes. Challenge: How to identify deepfake in real-time.	99%
2021	Gihun Lee and Mihui Kim(Lee and Kim 2021b)	FaceForensics++, DFDC	✗	✓	When compared to a CNN, MTCNN improves face detection accuracy. Limitation: quality of the dataset is weak and not diversified.	97%
2021	Yu-Cheng Liu et al.(Liu et al. 2020)	FaceForensic++ and Celeb-DF datasets	✗	✓	Different loss designs and data characteristics result in the greatest single-model results for models trained with Arcface SoftMax and SoftMax losses.	99%
2022	Young-Jin Heo et al(Heo, Yeo, and Kim 2022)	DFDC Celeb-DF (v2) Datasets	✗	✓	If just the spatial feature is considered, motion between neighboring synthetic pixel portions within a single frame may be missed	99%
2022	Aya Ismail et al.(Ismail et al. 2022)	Celeb-DF, Faceforencics++	✗	✓	YOLO (You only look once) technique used for detecting facial frames in videos	95.56%

2022	Suganthi ST et al.(Suganthi et al. 2022)	FFHQ, 100k-faces, DFFD, CASIA webface.	✗	✓	Deepfake image detection using fisher face with Local Binary Pattern Histogram (LBPH) and performance, accuracy, and sensitivity are higher than SVM, CNN, and KNN.	98.82%
2022	Goajian Wang et al.(Wang et al. 2021)	Celeb-DF(V2)	✗	✓	The only vector discriminative feature descriptions that is efficient and fast for deepfake detection is the Fused Facial Region Feature Descriptor (FFR-FD).	82.2%
2023	Sohail Ahmed Khan et al (Khan et al. 2022)	FaceForencics++ and DFDC	✗	✓	Describe a hybrid transformer network and an early fusion approach. This model employs two networks for feature extraction: XceptionNet and EfficientNet-B4	98.2%
2023	Abhishek Doshi et al (Doshi et al. 2022)	FaceForencics++	✗	✓	They suggested a system that employs transformers to extract spatiotemporal information, as well as a system that uses a video vision transformer to recognise any video in real time but without audio.	99%
2023	Yogewsh patel,Sudeep Tanwar et al(Patel et al. 2023).	CelebA, FFHQ, GDWCT, AttGAN, STARGAN, StyleGAN, StyleGAN2	✗	✓	Proposed a deep Convolution(D-CNN) model by applying it on 7 datasets for real and fake images. For face recognition Haar Cascade Algorithm is used and also compare with Mesonet and Meso inception net over CelebDF.	97%

5. Evaluation on Deepfake Detection Techniques

The majority of deepfake detection is accomplished by machine learning and deep learning strategies, which comprise a variety of models illustrated in fig.4. It is possible that models designed to identify modified images or videos cannot be used to identify fabricated audio, however few algorithms can identify multimodal deepfake. This review identified the AI based models that which existing model can be used for which type of deepfake media for the detection purpose. Evaluation of these models are shown in below table 5 & 6. This evaluation will provide a way to further researchers for the development of novel deepfake detection model.

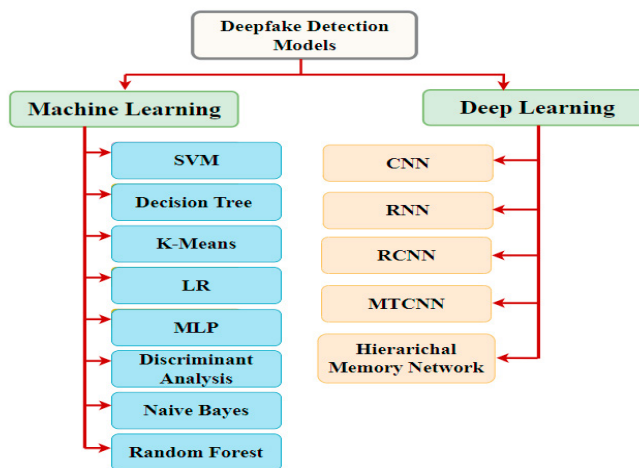


Fig. 4. Deepfake detection method based on ML and DL

Table 3. DL detection models used for deepfake detection

DL Techniques	Images	Video	Audio	Multimodal
Convolution neural network	✓	✓	✗	✗
Multi-task cascaded CNN	✓	✓	✗	✗
Recurrent neural network	✓	✓	✓	✓
Recurrent convolution neural network	✓	✓	✓	✓

Table 4. ML detection models used for deepfake detection

ML Techniques	Images	Video	Audio	Multimodal
Support Vector machine	✓	✓	✓	✓
K-Means	✓	×	×	×
Multilayer-perceptron	✓	×	×	×
Random-forest	✓	×	✓	×
Decision-Tree	✓	✓	×	×
Discriminant-Analysis	✓	×	×	×
Logistic-Regression	×	×	✓	×
Naive-Bayes	✓	×	×	×

On Evaluation of above these models in table.3 we have found that CNN is best for the images not for suitable the audio deepfakes detection and the RNN and RCNNs have been shown to be effective in detecting both audio and video-based deepfakes, as they can learn spatial and temporal patterns simultaneously. The table.4 shows the evaluation of ML deepfake detection models it is shown that the SVM is a ML algorithm which can detect all type of deepfakes because SVM uses Binary classification task for determining whether the given data is deepfake or not. The K-Means is a type of unsupervised ML algorithm that used to group similar data points together into cluster. It is not commonly used for deepfake detection but it can be used for image segmentation or for data compression.

6. Research Gaps and Future Perspective

Deepfake has been identified as a developing technology that brings numerous concerns

On various platforms, detection tools would be used to demonstrate the media's integrity. The scientific community has made the following observations on upcoming challenges and tasks based on the papers included in this survey.

- Existing methods are generally concerned with the front face, Face-swapping, or face re-enactment. Future deepfake algorithms may significantly change the upper or entire body.
- There is no apparent technique in existing deep learning detection methods for determining the number of layers required and which architecture is optimal for detection of deepfake.
- Existing datasets are not reliable enough to take into account different deepfake-generating techniques. Fewer data samples based on faked images have been seen. The reliability and variety of the datasets strongly impact how well detection techniques work. Future work might suggest the creation of a more comprehensive and diverse dataset.
- Since this technology progresses daily, it's possible that social media platforms will be misused. In order to stop the propagation of misinformation, it will be necessary to implement a deepfake detection method into these platforms. To effectively enforce this obligation, the legal obligation could be specified
- The current detection techniques focus on locating the flaws in deepfake-generating techniques and use them as classification artefacts. Future deep fakes may lack access to such knowledge or data, especially in an adversarial environment where attackers consciously strive to leave no evidence of their activity. So, the futuristic models must improve their robustness, generalizability, and scalability approaches.
- Collaboration between academia and industry can help to accelerate the development of new deepfake detection methods. This collaboration could lead to the creation of more sophisticated and effective detection tools

7. Conclusion and Future Scope

Deepfake generators and detectors are engaged in a constant arms race, in which each attempt to surpass the other by using more advanced and efficient methods. This paper discussed the benchmark deepfake image and video dataset, generation methods along with face manipulation techniques, detection techniques based on ML and DL, and research gaps with future perspectives. From the review analysis, it can be inferred that DL approaches are superior to ML because existing DL-based detection techniques have an accuracy of about 95% compared to ML techniques 86%. This study evaluated that among DL techniques, CNN models are mostly used for visual (image and video) deepfake detection whereas RNN or RCNN can be used for multimodal (image, video & audio) detection. In addition, the

SVM model is the only one of the machine learning techniques that can be used for multimodal deepfake detection. In future, we will use this study as tool for implementation purpose and propose a novel model to detect deepfake media and will overcome the research gaps of deepfake detection.

References

- Akhtar, Zahid, Dipankar Dasgupta, and Bonny Banerjee. 2019. "Face Authenticity: An Overview of Face Manipulation Generation, Detection and Recognition." *SSRN Electronic Journal*. doi: 10.2139/SSRN.3419272.
- Al-Dhabi, Yunes, and Shuang Zhang. 2021. "Deepfake Video Detection by Combining Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN)." in *2021 IEEE International Conference on Computer Science, Artificial Intelligence and Electronic Engineering, CSAIEE 2021*.
- Anon. n.d.-a. "Detecting and Simulating Artifacts InGANfake Images," in Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS), Dec. 2019, Pp. 1?6. - Google Search." Retrieved June 18, 2023 ([https://www.google.com/search?q=%60Detecting+and+simulating+artifacts+inGANfake+images%2C%27%27+in+Proc.+IEEE+Int.Workshop+Inf.+Forensics+Secur.+\(WIFS\)%2C+Dec.+2019%2C+pp.+1%156.&rlz=1C1ONGR_en__1033__1033&oq=%60Detecting+and+simulating+artifacts+inGANfake+images%2C%27%27+in+Proc.+IEEE+Int.Workshop+Inf.+Forensics+Secur.+\(WIFS\)%2C+Dec.+2019%2C+pp.+1%156.&aqs=cchrome..69i57.969j0j4&sourceid=chrome&ie=UTF-8](https://www.google.com/search?q=%60Detecting+and+simulating+artifacts+inGANfake+images%2C%27%27+in+Proc.+IEEE+Int.Workshop+Inf.+Forensics+Secur.+(WIFS)%2C+Dec.+2019%2C+pp.+1%156.&rlz=1C1ONGR_en__1033__1033&oq=%60Detecting+and+simulating+artifacts+inGANfake+images%2C%27%27+in+Proc.+IEEE+Int.Workshop+Inf.+Forensics+Secur.+(WIFS)%2C+Dec.+2019%2C+pp.+1%156.&aqs=cchrome..69i57.969j0j4&sourceid=chrome&ie=UTF-8)).
- Anon. n.d.-b. "https://www.kaggle.com/datasets/xhlulu/140k-real-and-fake-faces."
- Appel, Markus, and Fabian Prietzel. 2022. "The Detection of Political Deepfakes." *Journal of Computer-Mediated Communication* 27(4). doi: 10.1093/jcmc/zmac008.
- Arora, Bhavna. 2016. "Exploring and Analyzing Internet Crimes and Their Behaviours." *Perspectives in Science* 8:540–42. doi: 10.1016/J.PISC.2016.06.014.
- Brömme, A., C Busch, A. Dantcheva, C. Rathgeb, A. Uhl, Ali Khodabakhsh, Raghavendra Ramachandra, Kiran Raja, Pankaj Wasnik, and Christoph Busch. 2018. *Gesellschaft Für Informatik*.
- Doshi, Abhishek, Abhinav Venkatadri, Sayali Kulkarni, Vedant Athavale, Akhila Jagarlapudi, Shraddha Suratkar, and Faruk Kazi. 2022. "Realtime Deepfake Detection Using Video Vision Transformer." *IBSSC 2022 - IEEE Bombay Section Signature Conference*. doi: 10.1109/IBSSC56953.2022.10037344.
- F. Matern, C. Riess, and M. Stamminger, ``. n.d. "F. Matern, C. Riess, and M. Stamminger, ``Exploiting Visual Artifacts to Expose Deepfakes and Face Manipulations," in Proc. IEEE Winter Appl. Comput. Vis. Workshops (WACVW), Waikoloa Village, HI, USA, Jan. 2019, Pp. 8392, Doi: 10.1109/WACVW.2019.00020."
- Hangloo, Sakshini, and Bhavna Arora. 2021. "Fake News Detection Tools and Methods -- A Review."
- Hangloo, Sakshini, and Bhavna Arora. 2022. "Combating Multimodal Fake News on Social Media: Methods, Datasets, and Future Perspective." *Multimedia Systems* 28:6 28(6):2391–2422. doi: 10.1007/S00530-022-00966-Y.
- Heo, Young Jin, Woon Ha Yeo, and Byung Gyu Kim. 2022. "DeepFake Detection Algorithm Based on Improved Vision Transformer." *Applied Intelligence*. doi: 10.1007/s10489-022-03867-9.
- Huang, Jiajun, Xueyu Wang, Bo Du, Pei Du, and Antgroup Chang Xu. n.d. *DeepFake MNIST+: A DeepFake Facial Animation Dataset*.
- Ismail, Aya, Marwa Elpeltagy, Mervat S. Zaki, and Kamal Eldahshan. 2022. "An Integrated Spatiotemporal-Based Methodology for Deepfake Detection." *Neural Computing and Applications*. doi: 10.1007/s00521-022-07633-3.
- Karras, Tero, Samuli Laine, and Timo Aila. 2019. "A Style-Based Generator Architecture for Generative Adversarial Networks." Pp. 4396–4405 in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Vols. 2019-June. IEEE Computer Society.
- Khan, Sohail Ahmed, Mediafutures Bergen, Norway Duc, Tien Dang-Nguyen, and Duc-Tien Dang-Nguyen. 2022. "Hybrid Transformer Network for Deepfake Detection." *ACM International Conference Proceeding Series* 8–14. doi: 10.1145/3549555.3549588.
- Lee, Gihun, and Mihui Kim. 2021a. "Deepfake Detection Using the Rate of Change between Frames Based on Computer Vision." *Sensors* 21(21). doi: 10.3390/s21217367.
- Lee, Gihun, and Mihui Kim. 2021b. "Deepfake Detection Using the Rate of Change between Frames Based on Computer Vision." *Sensors* 21(21). doi: 10.3390/s21217367.
- Li, Yuezun, Xin Yang, Pu Sun, Honggang Qi, and Siwei Lyu. n.d. *Celeb-DF: A Large-Scale Challenging Dataset for*

DeepFake Forensics.

- Liu, Yu Cheng, Chia Ming Chang, I. Hsuan Chen, Yu Ru Ku, and Jun Cheng Chen. 2020. “An Experimental Evaluation of Recent Face Recognition Losses for Deepfake Detection.” Pp. 9827–34 in *Proceedings - International Conference on Pattern Recognition*. Institute of Electrical and Electronics Engineers Inc.
- Masood, Momina, Mariam Nawaz, Khalid Mahmood Malik, Ali Javed, Aun Irtaza, and Hafiz Malik. 2022a. “Deepfakes Generation and Detection: State-of-the-Art, Open Challenges, Countermeasures, and Way Forward.” *Applied Intelligence*. doi: 10.1007/s10489-022-03766-z.
- Masood, Momina, Mariam Nawaz, Khalid Mahmood Malik, Ali Javed, Aun Irtaza, and Hafiz Malik. 2022b. “Deepfakes Generation and Detection: State-of-the-Art, Open Challenges, Countermeasures, and Way Forward.” *Applied Intelligence*. doi: 10.1007/s10489-022-03766-z.
- Nadimpalli, Aakash Varma, and Ajita Rattani. 2022. “GBDF: Gender Balanced DeepFake Dataset Towards Fair DeepFake Detection.”
- Neha, and Bhavna Arora. 2023. “Deep Learning Based Model for Deepfake Image Detection: An Analytical Approach.” *3rd International Conference on Innovative Mechanisms for Industry Applications, ICIMIA 2023 - Proceedings* 1019–27. doi: 10.1109/ICIMIA60377.2023.10426561.
- Patel, Yogesh, Sudeep Tanwar, Pronaya Bhattacharya, Rajesh Gupta, Turki M. Alsuwian, Innocent Ewean Davison, and Thoko Zile F. Mazibuko. 2023. “An Improved Dense CNN Architecture for Deepfake Image Detection.” *IEEE Access*. doi: 10.1109/ACCESS.2023.3251417.
- Paul, Olympia A. n.d. *Deepfakes Generated by Generative Adversarial Networks*.
- Rahmouni N, et al. n.d. “Rahmouni N, et al . Distinguishing Computer Graphics from Natural Images Using Convolution Neural Networks. In: 2017 IEEE Workshop on Information Forensics and Security (WIFS). IEEE.”
- Robert Chesney and Danielle Keats Citron. n.d. “ Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. Democracy, and National Security, 107, 2018.”
- Rössler, Andreas, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, and Matthias Nießner. 2018. “FaceForensics: A Large-Scale Video Dataset for Forgery Detection in Human Faces.”
- Rössler, Andreas, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, and Matthias Nießner. 2019. “FaceForensics++: Learning to Detect Manipulated Facial Images.”
- Sandotra, Neha, and Bhavna Arora. 2023. “A Comprehensive Evaluation of Feature-Based AI Techniques for Deepfake Detection.” *Neural Computing and Applications* 2023 1–29. doi: 10.1007/S00521-023-09288-0.
- Shad, Hasin Shahed, Md Mashfiq Rizvee, Nishat Tasnim Roza, S. M. Ahsanul Hoq, Mohammad Monirujjaman Khan, Arjun Singh, Atef Zaguia, and Sami Bourouis. 2021. “Comparative Analysis of Deepfake Image Detection Method Using Convolutional Neural Network.” *Computational Intelligence and Neuroscience* 2021. doi: 10.1155/2021/3111676.
- Suganthi, S. T., Mohamed Uvaze Ahamed Ayoobkhan, V. Krishna Kumar, Nebojsa Bacanin, K. Venkatachalam, Hubálovský Štěpán, and Trojovský Pavel. 2022. “Deep Learning Model for Deep Fake Face Recognition and Detection.” *PeerJ Computer Science* 8. doi: 10.7717/PEERJ-CS.881.
- Tolosana, Ruben, Ruben Vera-Rodriguez, Julian Fierrez, Aythami Morales, and Javier Ortega-Garcia. 2020. “DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection.”
- Uçar, M. K. 2020. “Classification Performance-Based Feature Selection Algorithm for Machine Learning: P-Score.” *IRBM* 41(4). doi: 10.1016/j.irbm.2020.01.006.
- Wang, Gaojian, Qian Jiang, Xin Jin, and Xiaohui Cui. 2021. “FFR_FD: Effective and Fast Detection of DeepFakes Based on Feature Point Defects.” *Information Sciences* 596:472–88. doi: 10.1016/j.ins.2022.03.026.
- Zhang, Tao. 2022. “Deepfake Generation and Detection, a Survey.” *Multimedia Tools and Applications* 81(5):6259–76. doi: 10.1007/s11042-021-11733-y.