International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 16, No.1 (Aug-24)

http://dx.doi.org/10.12785/ijcds/160156

Exploring Deepfake Detection: Techniques, Datasets and Challenges

Preeti Rana[1] and Sandhya Bansal[2]

[1,2]*Computer Science and Engineering, Maharishi Markandeshwar Engineering College, Maharishi Markandeshwar (Deemed to be University) Mullana, Haryana*

Received 18 Dec. 2023, Revised 26 Apr. 2024, Accepted 01 May 2024, Published 01 Aug. 2024

Abstract: Deepfake detection is an active area of research due to extensive use of deepfake media for spreading false information, manipulate public opinion and cause harm to individuals. This paper presents a critical and systematic review of 84 articles for deepfake generation and detection. We review the current state-of-the-art techniques for deepfake detection techniques by grouping them into four different categories: deep learning-based techniques, traditional machine learning-based, artifacts analysis-based and biological signal-based methods, the datasets used for training and testing deepfake detection models. We also discuss the evaluation metrics used to measure the effectiveness of these methods and the challenges and future directions of deepfake detection research. Our findings suggest that deep learning models demonstrate superior accuracy compared to other methods and artifacts analysis-based methods shows greater potential in precision but there is still room for improvement in detecting more sophisticated and realistic deepfakes.

Keywords: Deepfakes, Deep Learning, Artifacts, Biological Signals, machine learning.

1. Introduction

Deepfakes are synthetic media that are designed to blend the target facial features onto a source face video making detection difficult [1], [2], [3]. These are categorized as head puppetry, face swapping, and lip syncing. In these categories source person's head, face swapping, and lip syncing respectively are used for generating convincing videos [4]. Face swapping and lip syncing are popular one and shown in Figure 1.

Figure 1. DeepFakes Containing Face Swapping and Lip Syncing [5]

Reddit's Deepfakes user raised concerns with AI-generated explicit content in 2017 [6]. While these technologies have primarily been used for legitimate purposes, such as entertainment and education as shown in Figure 2, malicious actors have also taken advantage of them for illegal or unethical activities as shown in Figure 3.

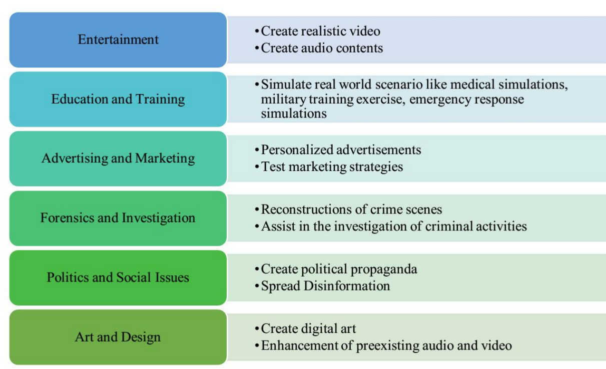

Figure 2. Applications of Deepfakes

A center for data innovation report by [7] found deepfakes contributed to 4% of social media misinformation during 2020 US election. Hence, a powerful deepfake detector is required to distinguish between true and fake information. Limited public awareness of deepfakes hinders detection by [8], limiting algorithm access to relevant data [9]. Technological advancements like GAN (Generative adversarial networks) by [10], DeepFaceLab by [11], Face-swap by [12] and Lensa AI by [13] etc. aids deepfake detection effectively. Some notable works focused on developing fea-

E-mail address: ranapreeti106@gmail.com, sandhya12bansal@gmail.com

TABLE I. Research Questions and Motivation

S. No	Research Questions	Motivation
RQ1	Which techniques are commonly used to detect deepfakes?	To demonstrate advancements in detecting Deepfake, categorization of these techniques and identify the challenges associated with existing detection methods.
RQ2	What are various datasets available for deepfake detection?	Having an up-to-date and accurate benchmark dataset, allowing comparison of deepfake detection algorithms.
RQ3	What are the various measures and metrics that can be utilized to determine the effectiveness of deepfake detection?	How to effectively compare and evaluate the deepfake detection algorithms and to find out the best algorithms evaluation is a necessity.
RQ4	What is the future scope of deepfake detection?	To discuss what areas, challenges have been researched and what still needs to be covered

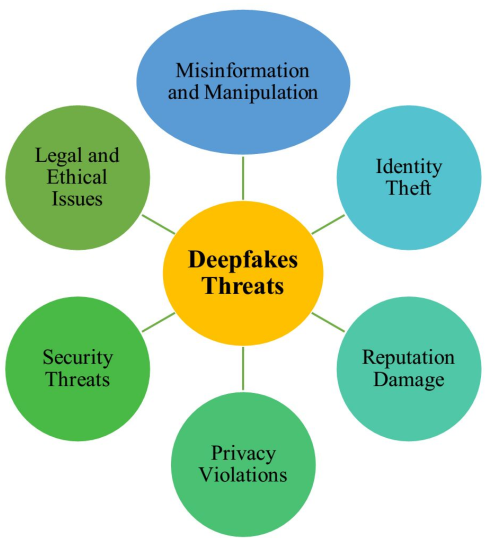

Figure 3. Deepfakes Threats

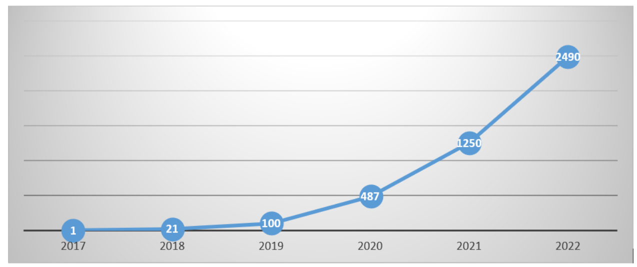

Figure 4. Exponential growth "deepfake detection" in Google Scholar since 2017

sible deepfake detection solutions include machine learning (ML) by [14], deep learning (DL) by [15], Frame difference analysis by [16], bio-signal analysis by [17] etc. Deepfake detection requires a combination of technical and human efforts, as well as continuous adaptation to the growing and changing landscape of deepfake technologies, which serves as a motivation behind this study and compilation of solutions in a single work.

The aim of this survey is to summarize the research progress regarding deepfake detection techniques as seen by the growth in paper published in Figure 4.

This survey will typically cover the deepfake detection models, dataset used, evaluation metrics, challenges and future directions. The major contribution of this survey paper is as follows:

1) It provides an updated and comprehensive overview of the various research works and methodologies proposed in the literature.
2) It analyses and categorizes these methodologies into different groups and evaluates their detection capabilities using different datasets. It can help in finding gaps in the existing techniques and hence providing a room for improvement.
3) The paper highlights the effectiveness of deep learning-based techniques in detecting deepfakes and can aid researchers in identifying potential research directions and areas for future exploration.

Remaining part of this survey is structured as follows. Section 2 outlines the research methodology employed for discovering and examining the available previous studies, along with the research questions and search standards. Next in section 3 a theoretical review of existing literature broadly in terms of detection approaches, dataset used and evaluation metrics has been provided. Then section 4 highlights the findings pertaining to in depth conducted survey in form of tables, pie charts and bar graphs. Section 5 outlines challenges and issues found during deepfake detection, followed by conclusion in section 6.

2. Research Methodology

The initial step in performing a survey by [18], is to identify and select the most relevant research papers that meet the inclusion criteria for the study. To accomplish this, a comprehensive search of the literature was conducted using renowned scientific databases. The survey paper has been focused on following research questions with motivation behind them as shown in Table I.

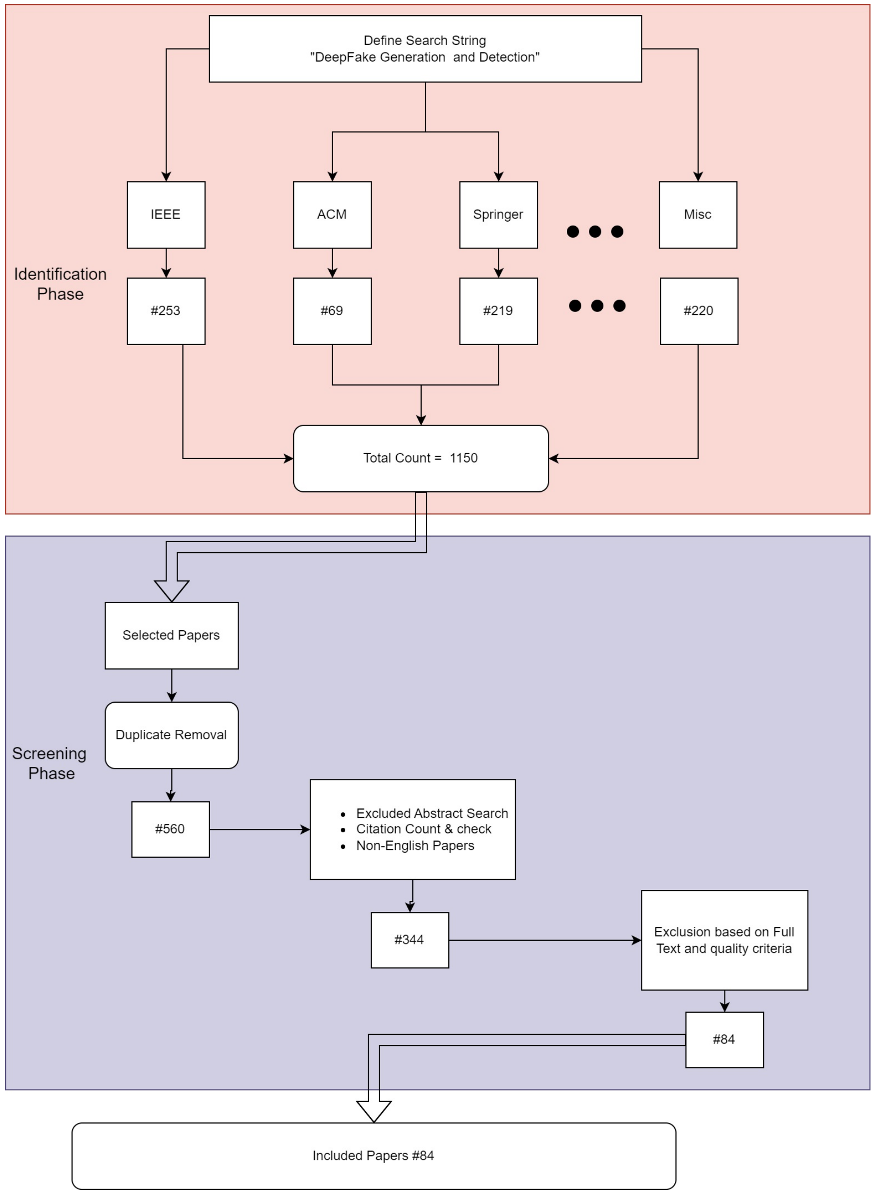

Figure 5. Research Article Screening Process

The methodology employed for search and selection of research articles is depicted in Figure 5.

Since the deepfake was started in 2018, we have set the article inclusion year from 2018 to 2023. We can clearly see from Figure 4 that very few papers were published in its initial years but during the year 2018 the research paced up. To identify the quality papers, first, the five most relevant databases, ACM, IEEE-Xplore, Web of Science, Science Direct, Springer, Google Scholar, and PubMed, were searched using the various keywords like "Deepfake Detection", "Deepfake tools OR methods OR techniques", "Deepfake Detection using deep learning OR machine learning OR biological analysis OR artifact analysis". A total of 1045 articles were fetched. The quality article finding process has been performed to ignore short articles, non-peer reviewed papers, book chapters and low-quality papers that were not able to give any technical information and scientific discussion. After this, a filter on the basis of title scan, removal of duplicates, magazine, conference proceedings, conference papers with pages less than 5, book chapters, and exclusion of review papers were applied and a total 560 papers were selected. Next, this count is reduced to 344 by applying a filtering process on the basis of citation count and quality check. Finally, selection criteria based on abstract scan and review questions were conducted that led to selection of final 84 articles for review.

3. Literature Review

A. Deepfake Detection Methods

Deepfake detection is the process of identifying and detecting artificially generated or manipulated media, such as images, videos, or audio, that have been created using DL techniques [19], [20]. Enhancing deepfake detection and mitigation is vital amid advancing technology. It involves analysing media properties like pixel values, frame rates, and using machine learning to identify fakery. Metadata analysis, such as device and location used, can also be useful. Various models are used for detecting Deepfake, these models are categorized in following groups: 1) Traditional Machine Learning (TML), 2) Deep Learning, 3) Biological Signals Analysis (BSB) and 4) Artifact Analysis. TML uses classical ML algorithms to detect the deepfakes, set of handcrafted features to aid machine learning algorithm. DL uses deep neural networks example Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs) whereas the BSB involves analysing biological signals such as facial expressions, eye movements, etc. Lastly the artifact analysis involves analysing the artifacts present in the video or image, such as compression artifacts, luminance changes, etc.

1) Traditional Machine Learning Based Methods

TML techniques like AdaBoost, SVM, and Random Forest excel in deepfake detection, offering resource efficiency, adaptability, and robustness to changing conditions, ideal for small datasets. Common models include Back Propagation Neural Networks, Decision Trees, Discriminant Analysis, K-Means clustering, Naive Bayes, Logistic Regression, and Multilayer Perceptron, applied widely in image and video analysis. [21] trained SVM classifier using feature points retrieved by one of many feature-point detectors, including FAST, KAZE, BRISK, ORB and HOG. A face recognition system is developed using VGG and NN that are sensitive to deepfake videos, with false acceptance rates of 85.62% and 95.00%, respectively [22]. Frequency domain analysis technique with a classifier is utilised to distinguish between genuine and fake photos, demonstrating promising performance in recognising deepfake images by [23]. Deepfake detection is accomplished by addressing the associated issue of attribution. Utilizing freely accessible FaceForensics++ datasets, authors show that training for attribution with a triplet-loss enhances generalisation while the performance with the same database decreases slightly by [24]. [25] analyses the functionality and operation of each unique algorithm utilising actual and fake facial recognition. They begin by normalising the images before doing an error level analysis using a SVM and the K-NN, which had an accuracy of 88.2% when compared to SVM's 86.8%. Additionally, variety of ways in which TML methods can be used for deepfake detection had shown by [26], [27], [28], [29].

2) Deep Learning Based Methods

DL models have emerged as a powerful tool in detecting deepfakes due to their ability to learn and identify complex patterns in images and videos automatically [30]. Pretrained models can also be fine-tuned for newer types of deepfake detection, reducing the volume of training data and computational resources required. Several types of DL models used are CNN, RNN and DenseNet. A CNN detection system with a compact architecture and an RNN to capture inconsistencies in face-swapping was presented in [31]. Tested on numerous deepfake videos from various sources, it achieved competitive results. Survey by [32] examines deepfake detection for distinguishing GAN-generated images. A robust, statistical approach aggregates features from various studies for classification. [33] conducted a study technology and concluded that ideal model for detection is SSTNet. [34] explores challenges in tracking AI-generated deepfakes. Experimental results highlight the necessary facial characteristics, spatial aggregations, and signal artifacts. Deepfake stack in [35] excels, achieving 99.65% accuracy. Evaluation of numerous deepfake schemes is performed by [36]. The study of [37] evaluated deepfake detectors on FaceForensics++ dataset, revealing up to 10.7% difference in error rate. A study on deepfake racial distribution found efficient training signals in "irregular" faces created by swapping faces. The study of [38] introduces vital deepfake detection components: Facial expression separator and classifier. It achieves 0.94-0.99 precision but reveals security flaws in adversarial scenarios. A new deep fake detection method (YOLO-CNN-XGBoost) is presented by [39] which works as a CNN Network perceptron at the highest possible level as well as achieves 90.73% correctness. In addition the study by [40], [41] also used DL to detect deepfakes.

3) Biological Signal based methods

BSB methods excel in deepfake detection, leveraging biological signals, real-time analysis, adaptability, and robustness, enhancing user experiences with natural solutions. One approach to identifying the generative model behind a deepfake and distinguishing deepfakes from real videos was presented in [42]. The researchers collected PPG data from actual and fraudulent videos and input them into a classification network, achieving 97.29% accuracy in identifying fake videos and 93.39% accuracy in identifying the source model. The study by [43], rPPG was introduced, which evaluates heart rate dynamics from facial videos. The study by [44] combination of deep and traditional motion magnification is developed that achieves 97.17% accuracy. The study by [45] AFMB-Net employs ML and heart rate analysis, offering deepfake detection with unforgeable heart rate, advancing GAN technology, even in low-quality videos.

4) Artifacts analysis- based methods

Artifacts in digital media, like facial deformities, are abnormalities introduced during processing or compression. In deepfakes, spotting unnatural facial movements, altered lighting, and skin tones helps detect artificial manipulation. Researchers enhance detection through methods like audio-video sync analysis, text examination, and biological signal analysis. Figure 6 shows facial deformities as facial artifacts.

A novel method for image fraud detection is introduced by leveraging dual-network recognition, achieving top results across various benchmarks (DFDC, Celeb-DF-v2, and FaceForensics++) in [46]. High quality fakes are challenging to detect visually thus spread false news [47]. Hence a study by He et al. [16] presents a novel fake detection method involving re-creation and enhanced generalization. Also, authors used artifact deepfake detection methods and showed better results and significant improvement over state of art work in [48], [49]. Artifacts-Disentangled Adversarial Learning (ADAL) presented by [50] obtain precise deepfake detection by separating the artifacts from unimportant data.

B. *Deepfake Datasets*

Deepfake datasets are collections of videos or images, including celebrities, politicians, and everyday people that are used for training and evaluating deepfake algorithms [51]. Datasets with insufficient samples are excluded from the as sample size is very important for deepfake detection as it is very easy to overfit the algorithm on the dataset. In order to provide a meaningful comparison, we prioritized datasets with a sufficient number of samples to ensure statistical significance and generalizability of the findings. Some of the popular deepfake datasets are:

1) UADFV: 1st generation dataset released in 2018, having real videos (49) and Deep Fake (49) videos with 294×500 pixels resolution, and average length of approximately 11.14 seconds [52].
2) Deepfake TIMIT: This comprises a set of real videos (320) and corresponding manipulated videos (640) of people speaking [53]. It is based on the TIMIT dataset, a commonly used speech recognition research dataset [54].
3) Celeb-DF v1 and v2: It contain celebrity face images with deepfake versions [55]. It includes over 590 real videos and 5639 deepfakes sourced from YouTube.
4) DeeperForensics-1.0: Large scale, highly diverse dataset with 60,000 videos, including 50,000 real and 10,000 fake samples, 18 million frames [56].
5) The DeepFake Detection (DFD): It consist of 363 real samples and 3068 fake videos [57]. The dataset was developed by paying actors, using open source deepfake generation methods.
6) Deepfake Detection Challenge (DFDC): This challenge is organized by Facebook and Partnership on AI, provides a dataset and platform for researchers to develop deepfake detection algorithms. It includes manipulated and real videos, with a leader board to track performance [58].
7) Face Forensics: It is the first and widely used benchmark dataset created by researchers from several institutions [59]. It includes videos from popular video datasets such as the YouTube Faces, the VoxCeleb [59].
8) Face Forensics in the Wild (FF-Waka FFW10K): It comprises of 10,000 reals as well as fake videos to reflect the Wild (real-world) scenarios [60].
9) KoDF: First Korean-language deepfake detection dataset created by researchers from several Korean institutions, depicting variety of subjects, containing Korean celebrities, politicians, as well as everyday people [61].
10) Video Forensics HQ: It contains only 45 persons contrast to another large dataset. Its aim was to answer the question of "how many persons are required to properly train a deepfake detector" [62].
11) FaceForensics++: It is a 2nd generation deepfake dataset and includes 1000 original video sequences altered with 4 automated face manipulation techniques: NeuralTextures, FaceSwap Face2Face, and Deepfakes [63]. The data was extracted from 977 videos of YouTube.

C. *Deepfake Evaluation Metrics*

This systematic literature review focuses on assessing the effectiveness of techniques for creating and detecting deepfakes analysing performance measures categorized as classifier evaluation and perceptual quality assessment.

1) *Deepfake Classifier Evaluation*

The Confusion Matrix is vital for assessing binary classifiers [64], summarizing effectiveness and identifying true/false positives/negatives. Table II shows a binary deepfake classifier confusion matrix with true positives/negatives and false positives/negatives.

TABLE II. Confusion Matrix for a Binary Classifier

	Deepfake (Estimated)	Real Video (Estimated)
Deepfake (Actual)	TRUE POSITIVES(TP)	FALSE NEGATIVES(FN)
Real Video (Actual)	FALSE POSITIVES(FP)	TRUE NEGATIVES(TN)

Another two key metrics are precision and recall. The percentage of samples that are truly positive among all the anticipated positives is the definition of precision for a classifier described as Equation 1.

$$PRECISION = \frac{TP}{TP + \alpha \times FP} \tag{1}$$

where $\alpha > 0$ is a weight determined by the ratio between the negative and positive samples.

Recall is the proportion of projected positive samples among the actual positive samples described as Equation 2.

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

To provide a more accurate assessment of the overall performance of a binary classifier, the most prevalent measures

Figure 6. Facial Artifacts (top row) and the Visual Artifacts (bottom row) [Güera, D et al., 18].

are accuracy and F-measure (F-score). Accuracy is the proportion of properly predicted samples (TP and TN) divided by the total categorised samples, as shown by Equation 3.

$$ACCURACY = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

The F-score is essentially a "Family of metrics", the most famous among all F-scores is the F1-score, which is the F-score with $f_i = 1$. Formally, it can be described as in Equation 4.

$$F_1 = \frac{PRECISION \times RECALL}{PRECISION + RECALL} = 2\times\frac{TP}{TP + FP + FN} \tag{4}$$

2) Deepfake Perceptual Quality Assessment Metrics

Assessing deepfakes quality requires subjective evaluation, commonly done through Perceptual Quality Assessment (PQA) methods [73]. For audio-visual signals is the Mean Opinion Score (MOS) [74]. A higher MOS score represents that the algorithm has a better ability to detect fake content and to preserve the original information. The MOS score is calculated by having evaluators rate the quality of the images or videos [75], and taking the average of the scores [76]. However, MOS is limited by requiring many evaluators, time, and may not reflect quality accurately. It also lacks details on errors/artifacts. These limitations make it important to use MOS in conjunction with other metrics, such as mean squared error (MSE) or structural similarity (SSIM) and peak signal-to-noise ratio (PSNR) [77], [78], [79].

MSE is determined by finding the average of the square of the differences between each pixel in the original image and the respective pixel in the reconstructed image as shown by Equation 5. A lower MSE value represents that the algorithm has a better ability to detect fake content and to preserve the original information.

$$MSE(X, Y) = \sum_{i=1}^{n}(y_i - x_i) \tag{5}$$

PSNR is calculated by comparing the maximum possible power of the original signal to the power of the difference between the original and reconstructed signals as shown in Equation 6. A higher PSNR value represents the algorithm has a better ability to detect fake content and to accurately preserve the original content.

$$PSNR = 10 \cdot log_{10}\left(\frac{R^2}{MSE}\right) \tag{6}$$

SSIM is a widely used image quality assessment metric that measures the structural similarity between the original and the synthesized image. It takes into account the luminance, contrast, and structure of the image. It is based on the idea that the human visual system is more sensitive to changes in structural information rather than changes in pixel values. The mathematical formulation of SSIM is as follows in Equation 7.

TABLE III. Comparison of top cited and prominent related woks on traditional ML, Deep learning, artifact analysis and bio-signal based methods

Paper	Type	Dataset Used	Accuracy (%)	Strength	Weakness
Guarnera et al., [19]	TML	CELEBA	90.22	Feature extraction via Convolution Traces.	Outdated dataset, Lower Accuracy.
Korshunov & Marcel, [22]	TML	VidTIMIT	91.03	Enhanced Performance	Lower Accuracy and High FAR.
Durall et al., [65]	TML	CELEBA FACE-HQ	91	Higher accuracy with less features.	Lower accuracy in comparison to SOTA method.
Rana et al., [26]	DL	FaceForensics++	97	High accuracy with less training time.	Testing on smaller dataset.
Wolter et al. [27]	DL	CelebA, FaceForensics++	96.5	Spatial conservation	-
Chen et al., [66]	DL	Celeb-DF v1 and v2	90.56	Less complex model	Lower Accuracy
Rana et al., [35]	DL	FaceForensics++	99	High Accuracy	Complex Model
De Lima et al., [67]	DL	Celeb-DFv2	98.26	High Accuracy	Time Extensive
Khormali &Yuan, [40]	DL	Celeb-DFv2	98.3	High Accuracy	Resource-intensive
L.Zhao et al., [41]	DL	CelebA	98.79	Adaptive convolutions, multi-feature fusion	Less robust
Matern et al., [68]	Artifact Analysis	Face2Face	86.6	Simple visual artifacts	Low Accuracy
Nguyen & Derakhshani, [20]	Artifact Analysis	Celeb-DF	88	Fast Training	Low Accuracy
Nirkin et al., [46]	Artifact Analysis	FaceForensics++, Celeb-DF-v2	96.98	High accuracy	Resource-intensive
Sun et al., [47]	Artifact Analysis	Faceforensics++	86.4	Orientation Invariance	Low Accuracy
He et al., [69]	BSB	CelebA	94.1	Robust Structure	Resource-intensive
Li et al., [50]	BSB	DFDC	98.7	High Accuracy	Resource-intensive
Dong et al., [48]	BSB	Celeb-DF	91.05	Effective Detector	Average Accuracy
Vinay et al., [45]	BSB	DeepFake TIMI	95.19	Use of skin color and heart beat analysis.	Low Accuracy
Elhassan et al., [70]	BSB	UADFV	96.47	Use of teeth and mouth movement.	Average Accuracy
Ciftci et al., [71]	BSB	CelebDF	91.50	Use of Biological signals	Low Accuracy
Jin et al., [72]	BSB	Face Forensics++	98	Luminance emphasis	Motion issues

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x + \sigma_y + c_2)} \quad (7)$$

where x and y are the two images being compared, μx and μy are the means of x and y, respectively, σx and σy are the standard deviations of x and y, respectively, σxy is the covariance between x and y, c_1 and c_2 are constants.

4. Discussion

This section will discuss findings for each research question based on the above mined literature review.

RQ1: Which techniques are commonly used to detect deepfakes?

Table III summarizes the comparison of most prominent papers of reviewed literature on on traditional ML, Deep learning, artifact analysis and bio-signal based methods. Table IV presents the comparative analysis of these four methods on various parameters. From these tables it can be concluded that DL and ML are effective but computationally expensive for deepfake detection. BSB based methods require specialized equipment and are currently less accurate. Artifact analysis-based methods are effective but may be less accurate on more sophisticated deepfake manipulations. However, hybrid methods (not shown in table V) combine different approaches to improve accuracy but can also be more computationally expensive. As every approach has its own strength and weakness, researchers could continue to explore the use of hybrid methods that combine different approaches to improve deepfake detection accuracy.

Figure 7 shows the categorization of commonly used deepfake detection methods. We had divided the deepfake detection model into five broad categories namely: DL , TML , hybrid, BSB and artifacts-based analysis based.

TABLE IV. Comparative Evaluation of Deepfake Detection Methods

Method	TML	DL	Artifact Analysis	BSB
Speed	Fast	Slow	Fast	Slow
Accuracy	Low	High	High	Mid
Complex inputs	No	Yes	Yes	Yes
Automated	No	Yes	No	No
Data Requirements	Small – Mid	High	Small	Small
Interpretability	High	Low	High	Low

RQ2: What are various datasets available for deepfake detection?

Table V shows a comparison of the datasets based on the number of real samples, deepfakes, source, generation, manipulation techniques, and year of publication.

FaceForensics++, a widely used deepfake research dataset, is valued for its size, diversity, and realism. UADFV and Celeb-DF are also common. Dataset popularity depends on specific research goals and data availability. Column generation signifies the data type, with first-generation being small and synthetic, second-generation being real-world and more challenging, and third-generation expected to be larger, more diverse, and complex, encompassing advanced deepfake techniques.

Figure 7 depicts the evolving deepfake datasets with more frames and identities, especially in DFDC. Generations show expanded data quality and variety, as in the third generation.

RQ3: What are the various measures and metrics that can be utilized to determine the effectiveness of deepfake detection?

Depending upon the metrics discussed in subsection 3-C1 a comparative Table VI has been made for four discussed approaches for deepfake detection.

From the Table VI it can be concluded that TML are simple and faster however very limited ability to handle complex images and videos with an average accuracy of 85%. DL methods on the other hand have this capability with average accuracy of 96%. Artifact-based methods outperforms others in terms of precision as shown in Figure 8. Accuracy of BSB method is higher than TML but lower than other two due to quality of the physiological sensors used to collect the signals and presence of noise. The winner of all methods where speed meets performance is artifact-based methods. Their accuracy is closer to DL.

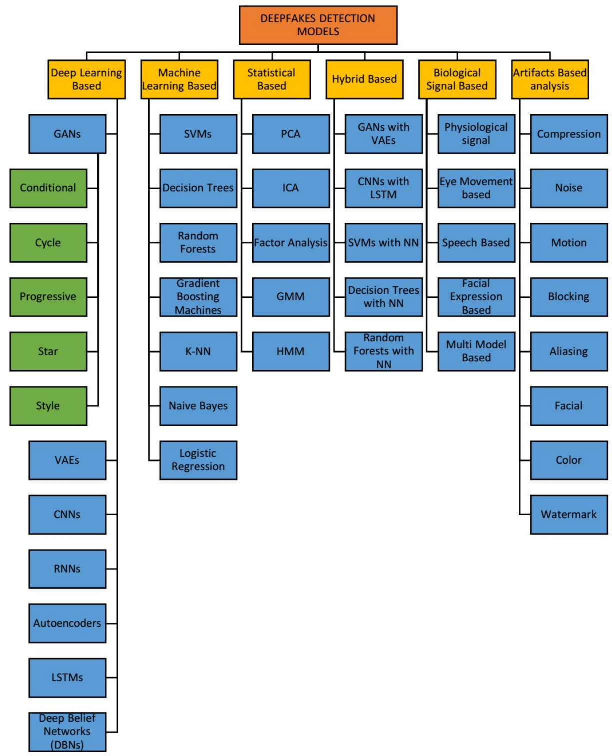

Figure 7. Deepfake Detection Models.

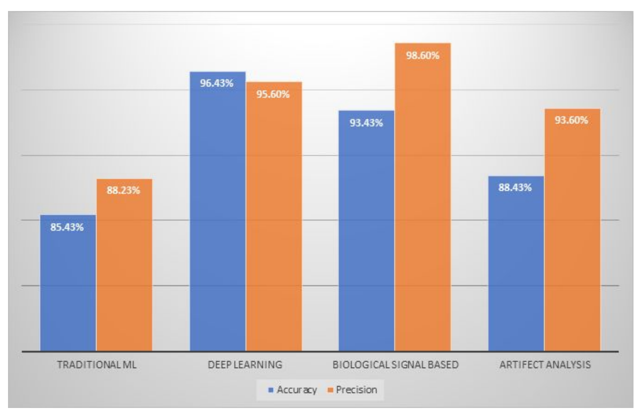

Figure 8. Average Accuracy and Precision for Deepfake Detection Methods

RQ4: What is the future scope of deepfake detection?

TABLE V. Deepfake Detection Dataset

Dataset	Year of Publication	Real Samples	Deepfakes	Source	Generation
DeepfakeTIMIT	2018	320	640	VidTIMIT	First
UADFV	2018	49	49	Existing	First
Celeb-DF v1 and v2	2019	590	5639	YouTube	First
DFD	2019	363	3,068	Paid Actors	Second
DFDC	2019	1131	4113	Volunteers	Third
FaceForensics++	2019	1,000	1,000	YouTube	Second
ForgeryNet	2019	99,630	1,21,617	Existing	Third
DeeperForensics-1.0	2020	50,000	10,000	Paid Actors	Third
FFIW10K	2020	10,000	10,000	YouTube	Second
KoDF	2021	62,166	1,75,776	Volunteers	First
VideoForensicsHQ	2021	261	1737	YouTube	First
Wild Deepfake	2021	3,805	3,509	Internet	First

TABLE VI. Metric Evaluation of Deepfake Detection Approaches

Category	# Papers	Metric	Min	Max	Avg	Std
TML	35	Accuracy	52.31%	95.68%	85.43%	11.26
		Precision	56.11%	97.48%	88.23%	15.06
		Recall	51.01%	94.38%	83.63%	12.56
		F1-Score	0.498	0.998	0.9268	13.04
DL	24	Accuracy	70.31%	100%	96.43%	7.26
		Precision	71.11%	100%	95.60%	13.06
		Recall	69.01%	100%	93.63%	7.56
		F1-Score	0.648	1.00	0.887	9.04
Artifact Analysis	16	Accuracy	81.31%	99.80%	93.43%	5.26
		Precision	82.11%	100.00%	98.60%	10.06
		Recall	82.01%	97.98%	97.63%	3.56
		F1-Score	0.7876	0.989	90.70%	5.04
BSB	9	Accuracy	69.31%	94.80%	88.43%	11.26
		Precision	66.11%	95.00%	93.60%	16.06
		Recall	67.01%	91.98%	91.63%	8.56
		F1-Score	63.76%	93.90%	84.70%	11.04

Deepfake detection is a rapidly evolving field, but there are still a number of challenges that need to be addressed in order to improve the accuracy and reliability of these approaches. Here are some of the key challenges:

1) Limited Availability of High-Quality and Diverse Datasets: Deepfake detection hindered by scarce high-quality, diverse datasets. Artifacts like splicing borders, low-quality faces impede algorithm effectiveness. Vital dataset enhancement needed [34].
2) Scalability: Researchers face a challenge with limited high-quality datasets. Scaling issues in current DL approaches hinder effective detection of Deep Fake techniques [33].
3) Poor Quality Datasets and Limited Real-World Relevance: Existing datasets for training deepfake detection lack real-world relevance due to poor visual quality. Performance on these may not translate to practical success. Limited diversity hampers algorithm efficiency [80].
4) Computational Optimization: DL struggles to keep up with escalating Deepfake quality. Algorithmic upgrades needed for effective recognition. Optimal layers and architecture uncertain [37].
5) Accuracy Optimization: The Deep Fake detection algorithm had a 65% accuracy and only identified 1/3rd of the Deep Fakes. 50% misclassification of real videos, 50% undetected, 35% false positives. Prioritize accuracy and efficiency.
6) Multiclass, cross-label, and localized recognition: Detecting Deep Fakes limited by binary categorization. Multiclass, cross-label, and localized recognition essential for detailed identification in complex scenarios.

The future of deepfake detection involves synergizing artifact-based and DL methods to overcome individual limitations. While artifact-based methods offer advantages, their dependency on specific creation processes necessitates integration with more versatile deep learning techniques.

Addressing dataset bias and enhancing robustness are crucial for improving overall effectiveness in countering advanced deepfake techniques.

5. Conclusion

This survey examines the current state-of-the-art methods for detecting deepfake, various datasets available for validation of methods, metrices for measuring performance of approaches. It also categorizes the existing approaches into 5 major groups. Additionally, it compares these categorized approaches in terms of various performance measuring parameters. The following summarization is provided:

- FaceForensics++ is one of the most popular datasets used for deep fake research.
- Detection accuracy is the most widely used performance metric.
- Experimental results show that DL techniques are effective in detecting Deepfake. DL models outperform non-DL models in terms of accuracy (96.34 %) whereas artifact analysis-based methods in terms of precision (98.60%).

References

[1] J. Kietzmann, L. W. Lee, I. P. McCarthy, and T. C. Kietzmann, "Deepfakes: Trick or treat?" *Business Horizons*, vol. 63, no. 2, pp. 135–146, 2020, https://doi.org/10.1016/j.bushor.2019.11.006.

[2] D. G. Johnson and N. Diakopoulos, "What to do about deepfakes," *Communications of the ACM*, vol. 64, no. 3, pp. 33–35, 2021, https://doi.org/10.1145/3447255.

[3] Y. Mirsky and W. Lee, "The creation and detection of deepfakes: A survey," *ACM computing surveys (CSUR)*, vol. 54, no. 1, pp. 1–41, 2021, https://doi.org/10.1145/3425780.

[4] F. Ding, G. Zhu, Y. Li, X. Zhang, P. K. Atrey, and S. Lyu, "Anti-forensics for face swapping videos via adversarial training," *IEEE Transactions on Multimedia*, vol. 24, pp. 3429–3441, 2021, https://doi.org/10.1109/TMM.2021.3098422.

[5] S. Lyu, "Deepfake detection," in *Multimedia Forensics*. Springer Singapore Singapore, 2022, pp. 313–331, https://doi.org/10.1007/978-981-16-7621-5_12.

[6] D. Gamage, P. Ghasiya, V. Bonagiri, M. E. Whiting, and K. Sasahara, "Are deepfakes concerning? analyzing conversations of deepfakes on reddit and exploring societal implications," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–19, https://doi.org/10.1145/3491102.3517446.

[7] K. Wahl-Jorgensen and M. Carlson, "Conjecturing fearful futures: journalistic discourses on deepfakes," *Journalism Practice*, vol. 15, no. 6, pp. 803–820, 2021, https://doi.org/10.1080/17512786.2021.1908838.

[8] A. Yadlin-Segal and Y. Oppenheim, "Whose dystopia is it anyway? deepfakes and social media regulation," *Convergence*, vol. 27, no. 1, pp. 36–51, 2021, https://doi.org/10.1177/13548565209239.

[9] N. C. Köbis, B. Doležalová, and I. Soraperra, "Fooled twice: People cannot detect deepfakes but think they can," *Iscience*, vol. 24, no. 11, 2021, https://doi.org/10.1016/j.isci.2021.103364.

[10] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE signal processing magazine*, vol. 35, no. 1, pp. 53–65, 2018, https://doi.org/10.1109/MSP.2017.2765202.

[11] I. Perov, D. Gao, N. Chervoniy, K. Liu, S. Marangonda, C. Umé, M. Dpfks, C. S. Facenheim, L. RP, J. Jiang *et al.*, "Deepfacelab: Integrated, flexible and extensible face-swapping framework," *arXiv preprint arXiv:2005.05535*, 2020, https://doi.org/10.48550/arXiv.2005.05535.

[12] D. Bitouk, N. Kumar, S. Dhillon, P. Belhumeur, and S. K. Nayar, "Face swapping: automatically replacing faces in photographs," in *ACM SIGGRAPH 2008 papers*, 2008, pp. 1–8, https://doi.org/10.1145/1399504.1360638.

[13] H. S. Sætra, "Generative ai: Here to stay, but for good?" *Technology in Society*, vol. 75, p. 102372, 2023, https://doi.org/10.1016/j.techsoc.2023.102372.

[14] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015, https://doi.org/10.1126/science.aaa8415.

[15] H. A. Khalil and S. A. Maged, "Deepfakes creation and detection using deep learning," in *2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*. IEEE, 2021, pp. 1–4, https://doi.org/10.1109/MIUCC52538.2021.9447642.

[16] J. Hu, X. Liao, J. Liang, W. Zhou, and Z. Qin, "Finfer: Frame inference-based deepfake detection for high-visual-quality videos," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 36, no. 1, 2022, pp. 951–959, https://doi.org/10.1609/aaai.v36i1.19978.

[17] B. Hosler, D. Salvi, A. Murray, F. Antonacci, P. Bestagini, S. Tubaro, and M. C. Stamm, "Do deepfakes feel emotions? a semantic approach to detecting deepfakes via emotional inconsistencies," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 1013–1022, https://doi.org/10.1109/CVPRW53098.2021.00112.

[18] E. Aromataris and A. Pearson, "The systematic review: an overview," *AJN The American Journal of Nursing*, vol. 114, no. 3, pp. 53–58, 2014, https://doi.org/10.1097/01.NAJ.0000444496.24228.2c.

[19] L. Guarnera, O. Giudice, and S. Battiato, "Deepfake detection by analyzing convolutional traces," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020, pp. 666–667, https://doi.org/10.1109/CVPRW50498.2020.00341.

[20] H. M. Nguyen and R. Derakhshani, "Eyebrow recognition for identifying deepfake videos," in *2020 international conference of the biometrics special interest group (BIOSIG)*. IEEE, 2020, pp. 1–5.

[21] F. F. Kharbat, T. Elamsy, A. Mahmoud, and R. Abdullah, "Image feature detectors for deepfake video detection," in *2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)*. IEEE, 2019, pp. 1–4, https://doi.org/10.1109/AICCSA47632.2019.9035360.

[22] P. Korshunov and S. Marcel, "Vulnerability assessment and detection of deepfake videos," in *2019 International Conference on Biometrics (ICB)*. IEEE, 2019, pp. 1–6, https://doi.org/10.1109/ICB45273.2019.8987375.

[23] H. Agarwal, A. Singh, and D. Rajeswari, "Deepfake detection using svm," in *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*. IEEE, 08 2021, pp. 1245–1249, https://doi.org/10.1109/ICESC51422.2021.9532627.

[24] A. Jain, P. Korshunov, and S. Marcel, "Improving generalization of deepfake detection by training for attribution," in *2021 IEEE 23rd International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2021, pp. 1–6, https://doi.org/10.1109/MMSP53017.2021.9733468.

[25] R. Rafique, M. Nawaz, H. Kibriya, and M. Masood, "Deepfake detection using error level analysis and deep learning," in *2021 4th International Conference on Computing & Information Sciences (ICCIS)*. IEEE, 2021, pp. 1–4, https://doi.org/10.1109/ICCIS54243.2021.9676375.

[26] M. S. Rana, B. Murali, and A. H. Sung, "Deepfake detection using machine learning algorithms," in *2021 10th International Congress on Advanced Applied Informatics (IIAI-AAI)*. IEEE, 2021, pp. 458–463, https://doi.org/10.1109/IIAI-AAI53430.2021.00079.

[27] M. Wolter, F. Blanke, R. Heese, and J. Garcke, "Wavelet-packets for deepfake image analysis and detection," *Machine Learning*, vol. 111, no. 11, pp. 4295–4327, 2021.

[28] J. Zhang, K. Cheng, G. Sovernigo, and X. Lin, "A heterogeneous feature ensemble learning based deepfake detection method," in *ICC 2022-IEEE International Conference on Communications*. IEEE, 2022, pp. 2084–2089, https://doi.org/10.1109/ICC45855.2022.9838630.

[29] S. Mathews, S. Trivedi, A. House, S. Povolny, and C. Fralick, "An explainable deepfake detection framework on a novel unconstrained dataset," *Complex & Intelligent Systems*, pp. 1–13, 2023, https://doi.org/10.1007/s40747-022-00956-7.

[30] A. M. Almars, "Deepfakes detection techniques using deep learning: a survey," *Journal of Computer and Communications*, vol. 9, no. 05, pp. 20–35, 2021, https://doi.org/10.4236/jcc.2021.95003.

[31] 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, 2018, pp. 1–6, https://doi.org/ 10.1109/AVSS.2018.8639163.

[32] S. Agarwal and L. R. Varshney, "Limits of deepfake detection: A robust estimation viewpoint," *arXiv preprint arXiv:1905.03493*, 2019, https://doi.org/10.48550/arXiv.1905.03493.

[33] R. Katarya and A. Lal, "A study on combating emerging threat of deepfake weaponization," in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*. IEEE, 2020, pp. 485–490, https://doi.org/10.1109/I-SMAC49090.2020.9243588.

[34] S. Lyu, "Deepfake detection: Current challenges and next steps," in *2020 IEEE international conference on multimedia & expo workshops (ICMEW)*. IEEE, 2020, pp. 1–6, https://doi.org/10.1109/ICMEW46912.2020.9105991.

[35] M. S. Rana and A. H. Sung, "Deepfakestack: A deep ensemble-based learning technique for deepfake detection," in *2020 7th IEEE international conference on cyber security and cloud computing (CSCloud)/2020 6th IEEE international conference on edge computing and scalable cloud (EdgeCom)*. IEEE, 2020, pp. 70–75, https://doi.org/10.1109/CSCloud-EdgeCom49738.2020.00021.

[36] Y. Nirkin, L. Wolf, Y. Keller, and T. Hassner, "Deepfake detection based on discrepancies between faces and their context," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 6111–6121, 2021, https://doi.org/10.1109/TPAMI.2021.3093446.

[37] L. Trinh and Y. Liu, "An examination of fairness of ai models for deepfake detection," *arXiv preprint arXiv:2105.00558*, 2021, https://doi.org/10.48550/arXiv.2105.00558.

[38] X. Cao and N. Z. Gong, "Understanding the security of deepfake detection," in *International Conference on Digital Forensics and Cyber Crime*. Springer, Cham, 2021, pp. 360–378, https://doi.org/10.1007/978-3-031-06365-7_22.

[39] A. Ismail, M. Elpeltagy, M. S. Zaki, and K. Eldahshan, "A new deep learning-based methodology for video deepfake detection using XG Boost," *Sensors*, vol. 21, no. 16, p. 5413, 2021, https://doi.org/10.3390/s21165413.

[40] A. Khormali and J.-S. Yuan, "Add: Attention-based deepfake detection approach," *Big Data and Cognitive Computing*, vol. 5, no. 4, p. 49, 2021, https://doi.org/10.3390/bdcc5040049.

[41] L. Zhao, M. Zhang, H. Ding, and X. Cui, "Mff-net: Deepfake detection network based on multi-feature fusion," *Entropy*, vol. 23, no. 12, p. 1692, 2021, https://doi.org/10.3390/e23121692.

[42] U. A. Ciftci, I. Demir, and L. Yin, "Fakecatcher: Detection of synthetic portrait videos using biological signals," *IEEE transactions on pattern analysis and machine intelligence*, 2020, https://doi.org/10.1109/TPAMI.2020.3009287.

[43] G. Boccignone, S. Bursic, V. Cuculo, A. D'Amelio, G. Grossi, R. Lanzarotti, and S. Patania, "Deepfakes have no heart: A simple rppg-based method to reveal fake videos," in *International Conference on Image Analysis and Processing*. Cham: Springer International Publishing, 2022, pp. 186–195, https://doi.org/10.1007/978-3-031-06430-2_16.

[44] U. A. Ciftci and I. Demir, "How do deepfakes move? motion magnification for deepfake source detection," *arXiv preprint arXiv:2212.14033*, 2022, https://doi.org/10.48550/arXiv.2212.14033.

[45] A. Vinay, N. Bhat, P. S. Khurana, V. Lakshminarayanan, V. Nagesh, S. Natarajan, and T. Sudarshan, "Afmb-net: Deepfake detection network using heart rate analysis," *Tehnički glasnik*, vol. 16, no. 4, pp. 503–508, 2022, https://doi.org/10.31803/tg-20220403080215.

[46] Y. Nirkin, L. Wolf, Y. Keller, and T. Hassner, "Deepfake detection based on the discrepancy between the face and its context," *arXiv preprint arXiv:2008.12262*, 2020, https://doi.org/10.48550/arXiv.2008.12262.

[47] X. Sun, B. Wu, and W. Chen, "Identifying invariant texture violation for robust deepfake detection," *arXiv preprint arXiv:2012.10580*, 2020, https://doi.org/10.48550/arXiv.2012.10580.

[48] S. Dong, J. Wang, R. Ji, J. Liang, H. Fan, and Z. Ge, "Towards a robust deepfake detector: Common artifact deepfake detection

model," *arXiv preprint arXiv:2210.14457*, vol. 10, 2022, https://doi.org/10.48550/arXiv.2210.14457.

[49] S. Dong, J. Wang, J. Liang, H. Fan, and R. Ji, "Explaining deepfake detection by analysing image matching," in *European Conference on Computer Vision*. Cham: Springer Nature Switzerland, 10 2022, pp. 18–35, https://doi.org/10.1007/978-3-031-19781-9_2.

[50] X. Li, R. Ni, P. Yang, Z. Fu, and Y. Zhao, "Artifacts-disentangled adversarial learning for deepfake detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 33, no. 4, pp. 1658–1670, 2022, https://doi.org/10.1109/TCSVT.2022.3217950.

[51] H. Zhao, W. Zhou, D. Chen, T. Wei, W. Zhang, and N. Yu, "Multi-attentional deepfake detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 2185–2194, https://doi.org/10.1109/CVPR46437.2021.00222.

[52] R. Tolosana, S. Romero-Tapiador, R. Vera-Rodriguez, E. Gonzalez-Sosa, and J. Fierrez, "Deepfakes detection across generations: Analysis of facial regions, fusion, and performance evaluation," *Engineering Applications of Artificial Intelligence*, vol. 110, p. 104673, 2022, https://doi.org/10.1016/j.engappai.2022.104673.

[53] S. A. Khan, A. Artusi, and H. Dai, "Adversarially robust deepfake media detection using fused convolutional neural network predictions," *arXiv preprint arXiv:2102.05950*, 2021, https://doi.org/10.48550/arXiv.2102.05950.

[54] T. Mittal, U. Bhattacharya, R. Chandra, A. Bera, and D. Manocha, "Emotions don't lie: An audio-visual deepfake detection method using affective cues," in *Proceedings of the 28th ACM international conference on multimedia*, 2020, pp. 2823–2832, https://doi.org/10.1145/3394171.3413570.

[55] 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 IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 3207–3216, https://doi.org/10.1109/CVPR42600.2020.00327.

[56] L. Jiang, R. Li, W. Wu, C. Qian, and C. C. Loy, "Deeperforensics-1.0: A large-scale dataset for real-world face forgery detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2889–2898, https://doi.org/10.48550/arXiv.2001.03024.

[57] N. Dufour and A. Gully, "Contributing data to deepfake detection research," *Google AI Blog*, vol. 1, no. 2, p. 3, 2019.

[58] B. Dolhansky, R. Howes, B. Pflaum, N. Baram, and C. C. Ferrer, "The deepfake detection challenge (dfdc) preview dataset," *arXiv preprint arXiv:1910.08854*, 2019, https://doi.org/10.48550/arXiv.1910.08854.

[59] A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "Faceforensics: A large-scale video dataset for forgery detection in human faces," *arXiv preprint arXiv:1803.09179*, 2018, https://doi.org/10.48550/arXiv.1803.09179.

[60] T. Zhou, W. Wang, Z. Liang, and J. Shen, "Face forensics in the wild," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 5778–5788, https://doi.org/10.1109/CVPR46437.2021.00572.

[61] 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*, 2021, pp. 10 744–10 753, https://doi.org/10.1109/ICCV48922.2021.01057.

[62] G. Fox, W. Liu, H. Kim, H.-P. Seidel, M. Elgharib, and C. Theobalt, "Videoforensicshq: Detecting high-quality manipulated face videos," in *2021 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2021, pp. 1–6, https://doi.org/10.1109/ICME51207.2021.9428101.

[63] 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*, 2019, pp. 1–11, https://doi.org/10.1109/ICCV.2019.00009.

[64] A. C. Lorena, A. C. De Carvalho, and J. M. Gama, "A review on the combination of binary classifiers in multiclass problems," *Artificial Intelligence Review*, vol. 30, pp. 19–37, 2008, https://doi.org/10.1007/s10462-009-9114-9.

[65] R. Durall, M. Keuper, F.-J. Pfreundt, and J. Keuper, "Unmasking deepfakes with simple features," *arXiv preprint arXiv:1911.00686*, 2019, https://doi.org/10.48550/arXiv.1911.00686.

[66] H.-S. Chen, M. Rouhsedaghat, H. Ghani, S. Hu, S. You, and C.-C. J. Kuo, "Defakehop: A light-weight high-performance deepfake detector," in *2021 IEEE International conference on Multimedia and Expo (ICME)*. IEEE, 2021, pp. 1–6, https://doi.org/10.1109/ICME51207.2021.9428361.

[67] O. De Lima, S. Franklin, S. Basu, B. Karwoski, and A. George, "Deepfake detection using spatiotemporal convolutional networks," *arXiv preprint arXiv:2006.14749*, 2020, https://doi.org/10.48550/arXiv.2006.14749.

[68] 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, 2019, pp. 83–92, https://doi.org/10.1109/WACVW.2019.00020.

[69] Y. He, N. Yu, M. Keuper, and M. Fritz, "Beyond the spectrum: Detecting deepfakes via re-synthesis," *arXiv preprint arXiv:2105.14376*, 2021, https://doi.org/10.48550/arXiv.2105.14376.

[70] A. Elhassan, M. Al-Fawa'reh, M. T. Jafar, M. Ababneh, and S. T. Jafar, "Dft-mf: Enhanced deepfake detection using mouth movement and transfer learning," *SoftwareX*, vol. 19, p. 101115, 2022, https://doi.org/10.1016/j.softx.2022.101115.

[71] U. A. Ciftci, I. Demir, and L. Yin, "How do the hearts of deep fakes beat? deep fake source detection via interpreting residuals with biological signals," in *2020 IEEE international joint conference on biometrics (IJCB)*. IEEE, 2020, pp. 1–10, https://doi.org/10.1109/IJCB48548.2020.9304909.

[72] X. Jin, D. Ye, and C. Chen, "Countering spoof: towards detecting deepfake with multidimensional biological signals," *Security and Communication Networks*, vol. 2021, pp. 1–8, 2021, https://doi.org/10.1155/2021/6626974.

[73] Y. Fang, H. Zhu, Y. Zeng, K. Ma, and Z. Wang, "Perceptual quality assessment of smartphone photography," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 3677–3686, https://doi.org/10.1109/CVPR42600.2020.00373.

[74] J. Wang, Z. Wu, W. Ouyang, X. Han, J. Chen, Y.-G. Jiang, and S.-N. Li, "M2tr: Multi-modal multi-scale transformers for deepfake detection," in *Proceedings of the 2022 international conference on multimedia retrieval*, 2022, pp. 615–623, https://doi.org/10.1145/3512527.3531415.

[75] R. C. Streijl, S. Winkler, and D. S. Hands, "Mean opinion score (mos) revisited: methods and applications, limitations and alternatives," *Multimedia Systems*, vol. 22, no. 2, pp. 213–227, 2016, https://doi.org/10.1007/s00530-014-0446-1.

[76] F. Ribeiro, D. Florêncio, C. Zhang, and M. Seltzer, "Crowdmos: An approach for crowdsourcing mean opinion score studies," in *2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2011, pp. 2416–2419, https://doi.org/10.1109/ICASSP.2011.5946971.

[77] K. Das, J. Jiang, and J. Rao, "Mean squared error of empirical predictor," *Ann. Statist*, vol. 32, no. 2, pp. 818–840, 2004, https://doi.org/10.1214/009053604000000201.

[78] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004, https://doi.org/10.1109/TIP.2003.819861.

[79] J. Korhonen and J. You, "Peak signal-to-noise ratio revisited: Is simple beautiful?" in *2012 Fourth International Workshop on Quality of Multimedia Experience*. IEEE, 2012, pp. 37–38, https://doi.org/10.1109/QoMEX.2012.6263880.

[80] N. Celebi, Q. Liu, and M. Karatoprak, "A survey of deep fake detection for trial courts," *arXiv preprint arXiv:2205.15792*, 2022, https://doi.org/10.48550/arXiv.2205.15792.

Preeti Rana is a Ph.D. Research Scholar at Maharishi Markandeswar (Deemed to be) University, Mullana,. She received her M. Tech degree from Maharishi Markandeswar (Deemed to be) University, Mullana, and B. Tech degree from Kurukshetra University(Haryana). Her areas of interest are Soft computing, Artificial Intelligence, and computer vision.

Dr. Sandhya Bansal is working as a Professor in the Department of Computer Science and Engineering at Maharishi Markandeswar (Deemed to be) University, Mullana, India. She has more than 18 years of teaching and research experience. She completed her PhD degree in Computer Science and Engineering in 2017. She is working as Co-PI in one of the funded projects by DST Haryana for Indian deaf and mute. She has published 4 patents and 45 articles including many in SCI, Scopus-indexed International Journals, and Conferences of repute. She has guided 01 Ph.D. student and has 05 registered students. Further, she is the editorial board member and reviewer of many journals of repute. Her current areas of research interest include soft computing, machine learning, vehicle routing problems, and computer vision techniques.