Er. Sunil Rathore, 2026, 14:6
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Deepfake Video Detection Through Facial Feature and Object-Level Analysis

Er. Sunil Rathore[1]
PG Student, Yagyavalkya Institute of Technology, Jaipur (Raj.), India[1].

Er. Laxmi Narayan Balai[2]
Assistant Professor, Yagyavalkya Institute of Technology, Jaipur (Raj.), India[2].

Abstract- The rapid development of generative artificial intelligence has made manipulated videos increasingly realistic and difficult to verify through visual inspection alone. Deepfake videos may imitate a person's identity, facial movements, expressions, or speech, creating risks for journalism, digital communication, privacy, and evidentiary media. This paper presents a feature-comparison framework for identifying suspicious temporal changes in video. The proposed workflow extracts frames from an input video, normalizes them, selects a reference frame, and compares subsequent frames using facial landmarks and object-level visual features. The principal indicators include facial landmark positions—especially around the eyes and mouth—colour or hue, contrast, and the relative arrangement of visible objects. Individual feature similarities are combined into an overall comparison score. A threshold such as 90 on a 0–100 scale may be used to flag frames for further inspection; it is a design parameter rather than a validated universal decision boundary. The framework is intended to support explainable, frame-level screening and localization of suspicious regions, not to replace expert verification. The source thesis identifies datasets such as FaceForensics++, DFDC, Celeb-DF, UADFV, and DeeperForensics-1.0 as potential evaluation resources. However, performance claims require reproducible testing on clearly defined train/test splits and diverse video conditions. The paper outlines the motivation, related approaches, proposed pipeline, evaluation plan, limitations, and future directions.

Keywords— deepfake detection, video forensics, facial landmarks, object features, temporal consistency, frame similarity.

I. INTRODUCTION

Digital video is widely used to communicate news, document events, support investigations, and preserve personal memories. Its value depends partly on the assumption that the content has not been misleadingly altered.

Deepfake methods challenge this assumption by using machine-learning models to synthesize or modify faces, expressions, and sometimes audio. A convincing video can make a person appear to say or do something that never occurred. The potential harms include misinformation, impersonation, reputational damage, privacy violations, and reduced public trust in authentic media. (Brundage et al., 2018; Botha & Pieterse, n.d.)

Deepfake creation systems commonly draw on generative adversarial networks (GANs), autoencoders, and other deep-learning architectures. In a GAN, a generator produces synthetic samples while a discriminator learns to distinguish generated samples from real ones. Repeated training can improve visual realism.

Autoencoder-based methods can learn compact representations and reconstruct facial content, while convolutional neural networks (CNNs) are widely used to learn spatial patterns from images. Sequence-oriented methods, including recurrent neural networks (RNNs) and long short-term memory (LSTM) networks, can model changes over time. (Goodfellow et al., 2014; Nguyen et al., 2020)

© 2026 Er. Sunil Rathore. his is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.



Fig. 1Deepfake and original Image

Detection remains difficult because artifacts vary across manipulation methods, video quality, lighting, compression, pose, and facial expression. A detector trained on one dataset may not generalize to a different source or a newly developed generation method. Some detectors also provide a prediction without clearly indicating which parts of a video influenced the decision. (Matern et al., 2019; Tolosana et al., 2020; Yu, 2021)

This paper focuses on a complementary, interpretable approach: compare selected facial and object features across frames and flag substantial or unusual deviations for investigation. The aim is to provide a structured screening method that can identify suspicious intervals and help a human reviewer examine the relevant frames. The proposed framework should be evaluated empirically before it is used for high-stakes decisions.

II. RESEARCH GAP, MOTIVATION AND OBJECTIVES

1. Research Gap

The source thesis identifies the need to compare original and potentially manipulated video content using measurable properties such as contrast, aspect ratio, frame resolution, colour information, and facial features. Many learning-based detectors rely on learned representations that can be effective but difficult to interpret. Hand-crafted feature approaches can be easier to explain, although they may be less robust to previously unseen manipulations. A practical framework should therefore balance interpretability, computational cost, and the ability to detect changes across a sequence.

A further challenge is distinguishing manipulation artifacts from natural variation. A person may move, blink, turn their head, or enter a different lighting condition. Camera motion, object movement, resizing, and video compression can also change frame-level measurements without indicating a deepfake. A simple threshold can consequently produce false alarms unless the method accounts for context and temporal continuity. (Tolosana et al., 2020; Yu, 2021)

2. Motivation

A frame-comparison workflow can help reviewers understand why a segment has been flagged. Instead of returning only a single video-level label, it can produce similarity scores and point to frames where the measured features differ from a selected reference. This is potentially useful for media triage, research demonstrations, and preliminary forensic review.

3. Objectives The objectives are

- Develop a structured pipeline for extracting and normalizing video frames.
- Represent facial and object-related information using measurable visual features.
- Compare each frame with a reference representation and calculate feature-level
- Identify frames or intervals that cross a configurable screening threshold.
- Define an evaluation plan using real and manipulated videos, with attention to false positives, false negatives, and generalization.
- Document limitations and ethical requirements for responsible use.

The central research question is: can interpretable comparisons of facial landmarks and object-level visual features help localize suspicious changes in a video sequence? This paper describes the proposed design; it does not claim validated accuracy without reproducible experimental evidence.

Datasets

To develop and train an algorithm to detect deepfakes, you need a large collection of real and fake films. Scholars usually use the following datasets to improve the understanding of deepfake detection:

- The Deepfake Detection Challenge, or DBDC, has the DFDC dataset, a crucial tool for deepfake detection. The video collection contains a wide range of artists, circumstances and modification techniques, with both real and modified content.
- The second dataset, FaceForensics++, is an augmentation of the original FaceForensics dataset. This dataset is great for testing deepfake detection algorithms, as it comprises deepfake videos created using multiple methods. These include techniques such as Face2Face, DeepDream and NeuralTextures. Celeb-DF is a dataset of celebrity deepfake videos. It is relevant to real-world events since it compiles deepfake videos with famous people.
- UADFV: The Unstructured Accessible Deepfake Video compilation contains both real and manipulated videos from different sources. The emphasis is on deepfake content, the more the better.
- DeeperForensics-1.0: This dataset is dedicated to source identification and facial modification operations. It contains both deepfake and real videos for comparative examination.
- NeuralTextures The NeuralTextures collection consists of deepfake films developed with the NeuralTextures technology. It offers a unique assortment of handpicked videos for academic research.
- Deepfake TIMIT: This dataset is a subset of TIMIT dataset that contains deepfake films. The TIMIT dataset is a famous dataset used in the research field of audiovisual deepfake detection.
- The FF++ dataset contains many manipulated movies, among which there are a few deepfakes created with the Face2Face, FaceSwap or Deepfakes methods. So detecting systems have many difficulties.
- WildDeepfake The WildDeepfake dataset contains deepfake videos that have been posted online, offering real-world instances of the spread of deepfake material. Google has created a data set to detect deepfakes with movies generated with the Face2Face, FaceSwap and NeuralTextures algorithms. There are videos that are "authentic" and unedited. (Dolhansky et al., 2020; Tolosana et al., 2020)

III. RELATED WORK

Research on deepfake detection spans visual artifact analysis, facial behaviour analysis, learned image representations, temporal modeling, and multimedia forensics. Early approaches often searched for visible inconsistencies, including irregular blending around a face, unnatural boundaries, colour mismatch, or imperfect rendering. These indicators can be useful in some cases, but generation quality and post-processing continue to improve. (Matern et al., 2019; Verdoliva, 2020; Nguyen et al., 2020; Yu, 2021)

CNN-based detectors learn spatial features from image regions and may identify patterns that are difficult to specify manually. Their performance can depend on the training distribution, dataset composition, compression, and manipulation type. Temporal approaches examine relationships between frames. RNN and LSTM models can represent sequential dependencies, while other methods compare motion, facial landmarks, or changes in expressions. Eye blinking and mouth movement have been investigated as useful signals, but they are not definitive evidence: real people vary in their behaviour, and modern synthetic media can reproduce many natural patterns. (Afchar et al., 2018; Nguyen et al., 2020; Almars, 2021)

Public research datasets offer different conditions and manipulation methods. The thesis discusses FaceForensics++, the Deepfake Detection Challenge (DFDC), Celeb-DF, UADFV, DeeperForensics-1.0, NeuralTextures, and Deepfake TIMIT. They differ in scale, source material, generation methods, and difficulty. Results across datasets should not be compared without considering evaluation protocol, video-level splits, preprocessing, and metrics. (Dolhansky et al., 2020; Tolosana et al., 2020)

The proposed framework differs from a purely end-to-end classifier by explicitly comparing a set of selected features against a reference frame. This supports a traceable score for each frame and may reduce the need for large model training when using simple features. Nevertheless, hand-designed similarity scores cannot be assumed to outperform learned detectors. A fair study should compare the proposed method with appropriate baselines and report both strengths and failure cases. The literature therefore motivates a hybrid future direction in which interpretable features complement learned representations rather than replacing them. (Mitra et al., 2021; Siegel et al., n.d.)

S. No.	Author(s) / Year	Paper name	Main focus and contribution	Limitations / research Gap
1	Goodfellow et al. (2014)	Generative Adversarial Networks (GANs)	Introduced the generator–discriminator training framework that became foundational to realistic synthetic image generation and later deepfake creation.	Foundational generation method; the supplied review does not provide a detection evaluation for this paper.
2	Li et al. (2019)	Deepfake detection: current challenges and next steps	Discusses deepfake applications, emerging problems, and the need for reliable detection systems.	The supplied review gives no specific benchmark results or detailed method description.
3	Suwajanakorn et al. (2017)	Synthesizing Obama: Learning Lip Sync from Audio	Demonstrates generating lip movements that synchronize a person's video with synthesized or supplied audio.	Highlights audio–visual manipulation; the supplied review does not state detection performance.
4	Matern et al. (2019)	Exploiting Visual Artifacts to Expose Deepfakes and Face Manipulations	Focuses on visual artifacts that can reveal manipulated faces and discusses the detection challenge as generation methods improve.	Artifact-based cues may be affected by changes in image quality, compression, or newer generation methods; these details are not quantified in the supplied text.
5	Brundage et al. (2018)	The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation	Considers the broader societal and security risks of malicious AI use, including potential misuse relevant to deepfakes, and mitigation approaches.	Broad risk and policy focus rather than a specific deepfake-video detector.
6	Wang et al. (2019)	Deepfake: A New Threat to Human Security	Provides a historical overview of deepfake development and the technology's implications for human security.	The supplied review does not identify a specific detection algorithm or quantitative evaluation.
7	Holt and Nery-Comacho (2020)	The Deepfake Dilemma: Technology and Policy Review	Discusses ethical concerns associated with digital manipulation and the importance of regulatory responses.	Policy-oriented discussion; no technical detection metrics are reported in the supplied review.
8	Li et al. (2020)	Deepfake Detection: A Survey	Reviews classical and deep-learning-based detection approaches, their challenges, and possible solutions.	The supplied review does not include detailed comparisons or complete bibliographic information.

S. No.	Author(s) / Year	Paper name	Main focus and contribution	Limitations / research Gap
9	Orlandi et al. (2020)	The Spread of Deepfake Videos in Online Social Networks	Examines how online social networks can distribute deepfake videos and how such distribution may support misinformation.	Social dissemination focus; the supplied review does not report detector-level performance.
10	Zheng et al. (year not specified)	Survey of Deepfake Text Generation	Reviews approaches for generating manipulated or synthetic text, along with possible applications and challenges.	The supplied review does not provide the full paper title, year, venue, or quantitative results.
11	Hsu and Jain (2020)	Media Forensics and DeepFakes: An Overview	Surveys media forensics and emphasizes verifying content integrity in response to deepfake technology.	Overview-level contribution; specific methods and performance values are not stated in the supplied text.
12	Authors not identified in supplied review [12]	Identification of DeepFakes and Face Manipulations	Discusses face manipulation, identity changes, facial-attribute edits, GAN fingerprints, generalization problems, and datasets such as UADFV and FaceForensics++.	The review highlights dataset dependence, compression sensitivity, limited public resources, and reduced generalization to unseen conditions.
13	Authors not identified in supplied review	Techniques for Multimedia Alteration (including DFAE, FSGAN, NTH, and StyleGAN)	Compares manipulation approaches and describes differences in temporal stability, lighting, skin-tone realism, head-pose handling, and visual quality.	The supplied review notes difficulties with low-light conditions, skin tones,

IV. PROPOSED FRAMEWORK

The proposed framework contains six stages: video ingestion, frame sampling, preprocessing, feature extraction, feature comparison, and decision support. (Mitra et al., 2021; Siegel et al., n.d.)

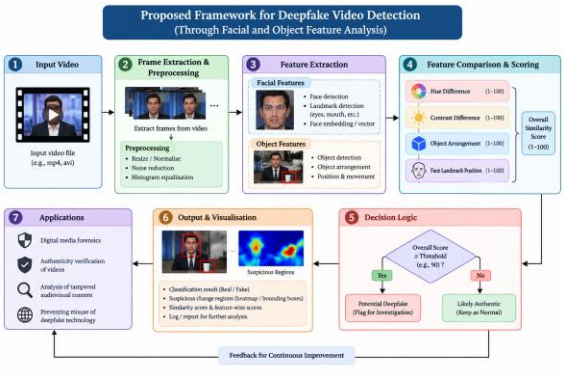


Fig2. Proposed framework of Facial and Object Visual Integrity Screening Framework

1. Video ingestion and frame sampling

The system receives a video file and reads its properties, including frame rate, duration, dimensions, and codec information where available. Frames may be sampled at a fixed interval or processed sequentially. The sampling policy must be recorded because sparse sampling can miss short-lived artifacts, whereas dense sampling increases computational cost.

2. Preprocessing

Each selected frame is converted into a consistent representation. Typical operations include resizing to a defined dimension, converting colour space where necessary, and checking that frames were decoded successfully. Original frame indices and timestamps should be retained so flagged frames can be located in the source video. Preprocessing must avoid introducing distortions that could be mistaken for manipulation.

3. Reference representation

A reference frame or reference feature set is selected. The first usable frame is a simple option, but it may be unsuitable if the face is initially occluded, blurred, or turned away. A more robust implementation can select a high-quality frame with a sufficiently visible face and stable objects. The selected reference and its selection criteria should be documented.

4. Feature extraction

The framework records facial landmarks, particularly around the eyes and mouth, alongside colour or hue statistics, contrast, and object arrangement. Facial landmark coordinates should be normalized relative to face size and position so that small camera movements do not dominate the score. Object features may describe the position, scale, or visual properties of detected regions. Missing or unreliable features should be marked as unavailable rather than assigned an arbitrary similarity value.

5. Feature comparison

For each subsequent frame, its features are compared with the reference features. The system calculates individual similarity values and combines available values into an aggregate score. The score should be designed so that higher values consistently mean greater similarity. Feature weights, normalization, and handling of missing values must be fixed before testing to avoid tuning the system to the evaluation set.

V. ALGORITHM AND DECISION LOGIC

Let the reference frame be F_0 and a sampled frame be F_i. The feature vector for frame i can be represented as $X_i = [L_i, C_i, H_i, O_i]$, where L represents facial-landmark information, C represents contrast, H represents colour or hue statistics, and O represents object-level information. The exact feature definitions depend on the implementation and should be reported in a reproducible experiment.

For each feature family k, compute a normalized similarity s_{ik} in the interval [0, 100], where 100 denotes maximum similarity under the selected metric. The aggregate similarity may be calculated as:

$$S_i = (\Sigma_k\, w_k\, s_{ik}) / (\Sigma_k\, w_k)$$

where w_k is the predefined weight for feature family k. Only valid feature families should contribute to the numerator and denominator. If a required feature cannot be measured reliably, the system should record the reason and avoid silently treating the missing value as a match.

A configurable threshold T can be used for screening. For example, if T = 90, frames with S_i below 90 may be flagged for review. This is only an illustrative starting point from the thesis concept; the threshold must be selected using validation data and may need to vary with the operating objective. A low similarity score does not prove that a video is fake. It indicates that the frame differs from the chosen reference according to the measured features.

High-level algorithm

- Read the input video and record metadata.
- Sample frames and preserve timestamps.
- Preprocess frames consistently.
- Select a reliable reference frame.
- Extract facial landmarks, colour, contrast, and object features.
- Compute feature-level similarity for every sampled frame.
- Combine valid feature scores into an aggregate similarity score.
- Flag frames below the validated threshold.
- Group adjacent flagged frames into candidate intervals.
- Export scores, timestamps, and a review report.

Temporal grouping can reduce isolated alerts, but it should not hide brief manipulations. The output should present both the frame-level evidence and any aggregated interval-level result.

Implementation

A practical prototype can be implemented in Python. OpenCV can be used for video decoding, frame extraction, resizing, and basic image operations. A facial-landmark detector is needed to locate the eyes and mouth; its source, version, and configuration should be documented. Object detection may be added where object-level comparisons are required.

NumPy can support feature-vector calculations, while pandas or CSV output can store per-frame results. Visualization tools can plot similarity scores against timestamps and display flagged frames for human review.

The implementation should separate the pipeline into small modules: video reader, preprocessor, face and object feature extractor, similarity calculator, decision module, and report generator.

This design improves testability and makes it easier to replace a detector or feature metric without rewriting the entire program. Every run should save configuration details, including frame-sampling interval, resizing dimensions, feature weights, threshold, and software versions.

A suitable output table includes video identifier, frame index, timestamp, feature availability, landmark similarity, colour similarity, contrast similarity, object similarity, aggregate score, and flag status.

This enables reviewers to inspect the basis of a decision instead of relying on an unexplained label. A visual timeline can show score changes and mark suspicious intervals.

Care is required when interpreting raw differences. Lighting changes may strongly affect colour and contrast. Head movement may change landmark coordinates, and camera motion may shift object locations. Normalizing landmark positions relative to the detected face and comparing local regions can reduce some nuisance variation. However, preprocessing cannot eliminate every source of error.

The prototype should handle corrupt videos, missing frames, videos without a visible face, multiple faces, and low-resolution footage. It should not crash or invent a score when feature extraction fails. Where the evidence is insufficient, the correct output is "inconclusive" or "manual review required," rather than an unsupported real/fake classification.

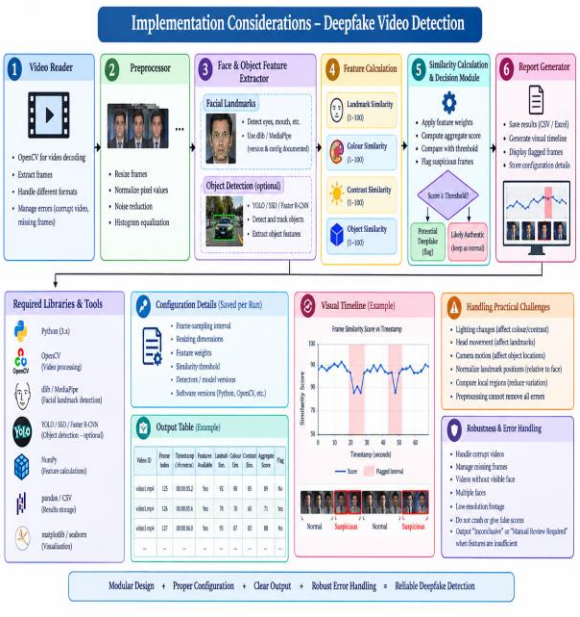


Fig3. Python-Based Implementation of the Proposed Deepfake Detection Framework

VI. EXPERIMENTAL EVALUATION

A rigorous evaluation is required before the framework can be described as accurate or robust. The thesis names several potential datasets, including FaceForensics++, DFDC, Celeb-DF, UADFV, and DeeperForensics-1.0. Dataset selection should reflect the intended use case and available permissions. Each dataset should be documented with its number of videos, authentic/manipulated distribution, manipulation types, and relevant quality conditions. (Dolhansky et al., 2020; Tolosana et al., 2020; Yu, 2021)The data should be split at the video or source-identity level to prevent near-duplicate frames from leaking across training, validation, and test partitions. If thresholds or feature weights are tuned, this must occur on validation data only. The final test set should be held out until the design is fixed. Cross-dataset testing is particularly important because performance may fall when a detector encounters a different generation method or compression profile. (Tolosana et al., 2020; Nguyen et al., 2020)Recommended metrics include accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (ROC-AUC), where appropriate. Confusion matrices should show true positives, true negatives, false positives, and

false negatives. Because the framework flags frames, evaluation should also report frame-level performance and video-level performance. If the intended output includes suspicious intervals, localization quality should be measured using a clearly defined interval or segment criterion. (Afchar et al., 2018; Almars, 2021; Yu, 2021)

Baselines may include a simple colour/contrast-only comparison, a facial-landmark-only comparison, an aggregate feature method, and a suitable established detector. An ablation study can determine whether each feature family adds value. Sensitivity analysis should test several thresholds rather than presenting one threshold as universally correct.

Results must be reported from actual runs with reproducible settings. No experimental accuracy is asserted in this paper because a complete, independently verifiable evaluation is not provided here. The evaluation plan is intended to make future claims testable and to expose cases where ordinary changes in pose, illumination, or compression cause false alarms.

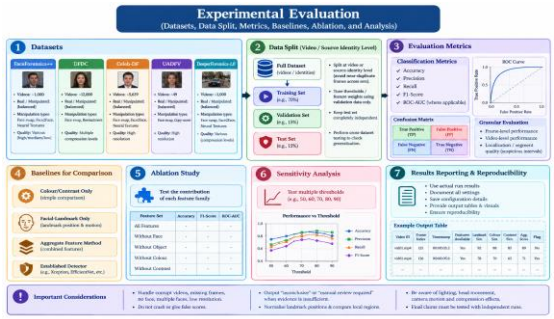


Fig. 4 Experimental Workflow for Deepfake Video Detection Using Facial and Object Feature Analysis

VII. CONCLUSION AND FUTURE WORK

This paper has presented an interpretable framework for screening deepfake videos through comparison of facial and object-related features across sampled frames. The workflow consists of video ingestion, preprocessing, reference-frame selection, feature extraction, similarity calculation, threshold-based flagging, and timestamped reporting. Facial landmarks around the eyes and mouth, colour or hue, contrast, and object arrangement provide complementary measurements that may help localize suspicious changes. (Verdoliva, 2020; Mitra et al., 2021)

The framework addresses an important practical need: detection systems should provide evidence that a human reviewer can inspect, not merely a binary label. Nevertheless, frame similarity is only an indicator of visual consistency. It cannot independently establish whether a video is authentic, and natural changes may produce alerts. The threshold and feature weights must be selected empirically, and the method should be compared with relevant baselines under reproducible conditions. (Tolosana et al., 2020; Siegel et al., n.d.)

Future work should prioritize robust reference-frame selection, face-size-normalized landmarks, motion compensation, scene-cut detection, and explicit handling of missing or unreliable features. A hybrid approach could combine interpretable features with CNN-derived representations and temporal models. Calibration and uncertainty estimation may help communicate when the evidence is weak. Evaluation should include cross-dataset tests, compression and resolution changes, different lighting conditions, multiple faces, and previously unseen manipulation methods.

The next practical step is to implement the pipeline, publish its configuration and evaluation protocol, and test it on authorized benchmark datasets. Results should include precision, recall, F1-score, ROC-AUC, confusion matrices, and examples of both successful detections and failure cases. Only after such testing can the framework's performance and practical value be established. The proposed method is best viewed as a transparent component of a broader media-verification process, supportingnot replacingexpert judgment and contextual investigation.

REFERENCES

1. Goodfellow, I., et al. (2014). Generative adversarial networks. arXiv preprint. [Complete

author list and arXiv identifier should be verified.]
2. Li, Y., et al. (2019). Deepfake detection: Current challenges and next steps. arXiv preprint. [Complete author list and arXiv identifier should be verified.]
3. Suwajanakorn, S., et al. (2017). Synthesizing Obama: Learning lip sync from audio. ACM Transactions on Graphics. [Volume, issue, pages, and DOI should be verified.]
4. Matern, F., et al. (2019). Exploiting visual artifacts to expose deepfakes and face manipulations. Proceedings of the IEEE International Conference on Computer Vision (ICCV). [Complete author list and page details should be verified.]
5. Brundage, M., et al. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. arXiv preprint. [Complete author list and arXiv identifier should be verified.]
6. Wang, L., et al. (2019). Deepfake: A new threat to human security. IEEE Internet Computing. [Volume, issue, pages, and DOI should be verified.]
7. Holt, T. J., & Nery-Comacho, J. (2020). The deepfake dilemma: A technology and policy review. Policy & Internet. [Author spelling and publication details should be verified against the original source.]
8. Li, Y., et al. (2020). Deepfake detection: A survey. arXiv preprint. [Complete author list and arXiv identifier should be verified.]
9. Orlandi, F., et al. (2020). The spread of deepfake videos in online social networks. IEEE Transactions on Information Forensics and Security. [Complete author list and publication details should be verified.]
10. Zheng, X., et al. (2020). Deepfake text: A survey. arXiv preprint. [Complete author list and arXiv identifier should be verified.]
11. Hsu, A., & Jain, A. (2020). Media forensics and deepfakes: An overview. IEEE Signal Processing Magazine. [Author initials and publication details should be verified.]
12. Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). DeepFakes and beyond: A survey of face manipulation and fake detection. arXiv:2001.00179.
13. Verdoliva, L. (2020). Media forensics and deepfakes: An overview. arXiv:2001.06564.
14. Dolhansky, B., Bitton, J., Pflaum, B., Lu, J., Howes, R., Wang, M., & Canton Ferrer, C. (2020). The deepfake detection challenge (DFDC) dataset. arXiv:2006.07397.
15. Afchar, D., et al. (2018). MesoNet: A compact facial video forgery detection network. Proceedings of the IEEE International Workshop on Information Forensics and Security. https://doi.org/10.1109/WIFS.2018.8630761
16. Yu, P. (2021). A survey on deepfake video detection. [Journal/publication details should be verified]. https://doi.org/10.1049/bme2.12031
17. Sohrawardi, S. J. (2020). DeFaking deepfakes: Understanding journalists' needs for deepfake detection. [Conference/publication details should be verified]. https://bpb-us-w2.wpmucdn.com/sites.northeastern.edu/dist/d/53/files/2019/11/CJ_2020_paper_64.pdf
18. Nguyen, T. T., et al. (2020). Deep learning for deepfakes creation and detection: A survey. arXiv:1909.11573.
19. Almars, A. M. (2021). Deepfakes detection techniques using deep learning: A survey. Journal of Computer and Communications. https://doi.org/10.4236/jcc.2021.95003
20. Wubet, W. M. (2020). The deepfake challenges and deepfake video detection. International Journal of Innovative Technology and Exploring Engineering, 9(6). [Page range and DOI should be verified.]
21. Botha, J., & Pieterse, H. (n.d.). Fake news and deepfakes: A dangerous threat for 21st century information security. [Publication details should be verified]. https://www.proquest.com/openview/f467fc8268b82bf3d2f2c3920490419e/1
22. Siegel, D., Kraetzer, C., Seidlitz, S., & Dittmann, J. (2021). Media forensics considerations on deepfake detection with hand-crafted features. Journal of Imaging, 7(7), 108. https://doi.org/10.3390/jimaging7070108
23. Mitra, A., Mohanty, S. P., Corcoran, P., & Kougianos, E. (2021). A machine learning based approach for deepfake detection in social media

through key video frame extraction. SN Computer Science. https://doi.org/10.1007/s42979-021-00495-x