IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

DEEPTRACE - Deepfake Detection & Source Extraction

B Malathi[1], M Manoj[2], P Vinayak[3], T Yashwanth[4]
[234]Department of Computer Science and Engineering,
[1]Associate Professor, Department of Computer Science and Engineering
Sreenidhi Institute of Science and Technology, Hyderabad, TS, India
[1]malathi.b@sreenidhi.edu.in, [2]hemasaimanojreddy7575@gmail.com, [3]vinayakapalvai@gmail.com,
[4]yashwanthtalapareddy123@gmail.com

Abstract: *At a time when AI-created media is rapidly evolving, distinguishing what is true from the manipulated footage has become an issue of paramount importance. Fake videos created through the process of deepfake can severely damage the reputation of the media industry as well as contribute to misinformation online. DeepTrace is an innovative deep learning based solution for the detection of fake videos. While other solutions may simply help identify videos as manipulated or not, DeepTrace detects whether the footage was manipulated while at the same time allowing for tracking of its origin, whether it is DeepFaceLab, FaceSwap or GAN-based manipulation. By means of employing advanced algorithms and state-of-the-art computer vision technologies, this service is able to detect subtle cues that suggest whether the video was faked by looking for inconsistencies in motion, facial patterns or texture which are not perceivable by humans. The system offers an easy-to-use web platform where the user can input video footage and get instant analysis results without requiring any special skills or knowledge. In addition to being a valuable service for detection of fakes, DeepTrace also helps restore accountability and reliability of digital media.*

Keywords: Deepfake Detection, Convolutional Neural Network, Computer Vision, Real-Time Analysis, Video Forensics, Source Tracing, Digital Media Integrity, Deep Learning.

I. INTRODUCTION

In recent times, the fast proliferation of the deepfake technology has made it harder to tell the difference between authentic video content and that which has been manipulated through various means. While the technology has its merits in certain applications such as entertainment and media, there are several negative implications associated with its use. It is therefore necessary to develop a robust system that analyzes the technology.

The DeepTrace algorithm is designed to tackle this problem by utilizing sophisticated deep learning algorithms for the purpose of recognizing deepfake videos. In particular, the technique pays attention not only to features present inside individual images but also to patterns emerging from frame sequences.

It operates by first performing frame extraction from the video in question before passing them through a convolutional neural network for analysis of visual discrepancies. This is followed by temporal analysis through which movement inconsistencies are established and analyzed. The results are then combined to give the ultimate conclusion about whether the video is genuine or artificial.

Apart from detecting the presence of deepfakes, there is also an ability to track back to the origin of the tampered content. Users can conveniently submit their videos via a straightforward and easy-to-use website interface.

Overall, DeepTrace hopes to restore confidence in online media through its efficient, accurate, and user-friendly deepfake detection technology.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

II. RELATED WORK

The process of detecting deepfakes has undergone substantial development due to the exploration of various approaches by researchers in the light of ever-changing complexities in synthetic media. Initially, most researches used Convolutional Neural Networks (CNNs) to detect visual anomalies in image-based deepfakes. The CNNs were designed to analyze anomalies within images including lighting problems, unnatural texture of the skin, and facial blending issues. Techniques such as MesoNet were quite efficient in terms of detecting any manipulation done to images and had a great advantage of rapid processing. However, the biggest flaw of these approaches is that they only analyzed individual frames without taking into account the relationships among frames.

Frequency-domain analysis techniques were incorporated into the process to enhance detection. Rather than analyzing the images themselves, researchers would first convert images from the spatial domain to the frequency domain through the use of techniques like the Fourier transform. Through this process, researchers could then detect invisible signs of tampering, such as inconsistent image compression and geometrical distortions. Although frequency-domain analysis yielded positive outcomes, especially in regards to early versions of deepfakes, this was no longer effective as new-generation technologies developed.

One of the key directions studied in relation to the research was the use of handcrafted features. In the early techniques used for forensics, people would manually design features that can help differentiate between authentic and tampered images, including noise, color, and compression features. However, when applied in real-world situations, such as where videos were being compressed and filtered, these handcrafted features did not work as well as expected.

With improvements in deepfake technology, scientists found that the analysis of temporal data was critical. As a result, sequence models like RNNs and LSTM were applied. Sequence models analyze the evolution of frames and can detect anomalies in terms of irregular blinking, unusual facial expressions, and lip synchronization issues. Sequence models made it easier to detect deepfakes since they are capable of recognizing behaviors that are not possible to identify using frame-based methods. However,

Some more recent work is focused on using multimodal systems that use video, speech, and context. For instance, one can evaluate whether the mouth movement is aligned with the speech audio or whether there are any mismatches between facial expression and vocal tone. Transformers and attention-based frameworks are other innovations that aim to leverage long-term dependencies in video sequences. Such systems are highly accurate but require more resources and effort to be deployed.

Even with all these developments, there are still some problems. First, many methods fail to generalize when it comes to detecting new deepfakes, as they only learn from one particular dataset. Second, the problem of biased datasets, lack of diversity, or changes in lighting or cameras affects the efficiency of the model. Third, the computational complexity involved is very high, thus making it hard to implement sophisticated methods on resource-limited hardware.

However, despite all these advancements, there are still various obstacles that need to be overcome in order to improve detection of deep fakes. The first obstacle is the problem of generalization – models tend to show high accuracy on previously trained samples, while having difficulties when it comes to detection of new types of deep fakes. Another problem is that of data bias – most of the datasets used in detection lack diversity, which means that no consideration is made regarding lighting conditions, ethnicity, and camera type.

Hybrid approaches, which attempt to leverage the benefits of both spatial and temporal approaches, were developed in order to improve performance further. Hybrid approaches usually employ convolutional neural networks for capturing spatial characteristics of each image frame, while using Long Short Term Memory or some other architecture for capturing temporal characteristics. As a result, hybrid approaches demonstrate improved performance compared to approaches relying on only one type of information. Indeed, numerous research studies prove the superior performance of hybrid approaches, particularly for deepfakes of high complexity and quality.

Apart from enhancing the level of accuracy, efforts have been made to develop a detectable system that provides justification for the output prediction by users. Deep learning-based detectors have been known to generate output

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

without clear explanation to justify their results. In order to solve this challenge, techniques such as Grad- CAM and attention models were proposed to show parts of the frame which contribute to the decision of the model.

Some sophisticated techniques have been developed that try to detect deepfake videos based on biological indicators and behavior patterns. This includes eye-blink rate analysis, analysis of heart rate signals using the skin tone change method, and micro-expression detection. Such biological indicators are hard to simulate, hence providing an added dimension of security for detecting deepfakes. However, these approaches depend heavily on external factors such as light intensity and camera resolution.

In summary, the area of detecting deepfakes has progressed from initial image-only techniques to more sophisticated methods such as hybrids and multi-modal. These techniques have all provided unique insights; however, no method alone provides complete protection against deepfakes. This has resulted in the creation of an integrated solution, which consists of different techniques being used together in order to create more reliable detection. The DeepTrace technique takes advantage of these improvements by integrating various methods while also adding important aspects such as explainability and source tracing.

III. METHODOLOGY

The methodology used by DeepTrace to detect deepfake videos includes a set of procedures that use both spatial and temporal features of the videos. The procedure entails preprocessing of the video file, extraction of the frames, analysis of the visual features via CNNs, and modeling of the sequences via LSTM. The method provides better accuracy in predicting whether a video is authentic or a deepfake due to its consideration of both visual and motion aspects.

Step 1: Video Upload

Step one involves the upload of a video file through the website interface. This system supports video files in popular formats such as MP4. An identifier is assigned to each uploaded video, and it acts as the video's job ID.

Step 2: Input Validation

This video is validated for correct formatting and size. This guarantees that only legitimate and executable videos reach the subsequent stages, thus preventing any errors from occurring during the execution process..

Step 3: Video Loading

The preprocessed dataset is divided into three subsets: Training set for model learning Validation set for tuning hyperparameters Testing set for evaluating final model performance This ensures unbiased evaluation and reliable accuracy measurement.

Step 4: Frame Extraction

The frames will be sampled from the video sequence at a constant rate, for instance, 10 frames per second. This technique allows capturing all the necessary details while balancing accuracy and efficiency.

Step 5: Frame Preprocessing

Every frame undergoes resizing to a standard resolution (224 × 224 pixels) and is represented using the RGB color model. This uniformity is important for all inputs and adheres to the demands of pre-trained deep neural networks.

Step 6: Data Normalisation

The pixel values are normalized using standardized values of mean and standard deviation (such as normalization in ImageNet). This helps to minimize any variance that may result from lighting differences.

Step 7: Feature Extraction using CNN

Each frame undergoes processing by means of deep CNN models such as ResNet-50. Hierarchical features are extracted ranging from simple edges to complicated facial structures. It allows the recognition of slight visual anomalies such as texture discontinuities, contour distortions, and unnatural blendings..

Step 8: Frame Level Prediction

The CNN classifies the image into probabilities for being real or fake images. This helps with localizing the detection and also detecting the frames that show more signs of being manipulated.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

Step 9: Sequence Formation
Frame-level features are grouped into sequences (e.g., sliding windows of 16 frames). This preserves the temporal order and allows the system to analyze motion continuity across frames.
Step 10: Temporal Analysis using LSTM
The LSTM model processes these sequences to capture temporal dependencies. It identifies anomalies such as irregular blinking, inconsistent head movement, and lip- sync mismatches, which are common in deepfake videos.
Step 11: Feature Fusion and Classification
After being fused together (feature fusion), their outputs are processed using fully connected layers. The output is a probability value that will be compared against a predefined threshold value (like 0.5) to determine whether the video is Real or Fake.
Step 12: Result display amd Visualization
Th The system presents the result through the interface, including classification, confidence score, and visual explanations like Grad-CAM heatmaps. These visual cues highlight suspicious regions and improve interpretability for users.

3.1. Algorithm:

The system takes a video file as input, which is uploaded by the user through a simple web interface in formats like MP4. This video is processed using trained deep learning models along with predefined parameters such as frame rate and threshold values. The system analyzes the video by extracting frames, preprocessing them, and passing them through the detection pipeline. After processing, the system provides an output indicating whether the video is Real or Fake, along with a confidence score. It may also include visual explanations like highlighted regions to show suspicious areas, and in some cases, provide possible source information for further investigation.

3.1.1. Step-by-Step Algorithm:

• Initiate the processing procedure.
• Obtain the input video by the user.
• Check the validity of the input video (format, size, integrity).
• Import the video and convert it to frames using a constant frame rate
• Perform preprocessing (resizing, normalization) on each frame
• Process the frame using CNN to extract features and classify
• Generate sequences out of the features
• Use LSTM to analyze the sequence
• Combine CNN and LSTM results to generate the final score.
• Output “Real Video” or “Fake Video”.

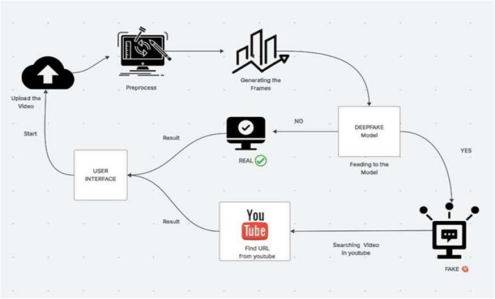
3.1.2. Feature Extraction

Feature extraction plays a critical role in the DeepTrace framework in which essential information is extracted from each frame of the videos to determine whether they are genuine or manipulated. Instead of analyzing the images in their raw form, the framework focuses on extracting patterns from the images like texture, edge detection, and facial features.
During this process, each image goes through the Convolutional Neural Network (CNN), which can be taken as an example like the ResNet-50 network. CNN learns on its own about the different kinds of features that need to be extracted from the images. Initially, in the early layers, CNN detects primitive features like edges and colors. But deeper into the layers, CNN recognizes complicated features like facial shape, skin tone, etc.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

The resultant features from the process above are then turned to vector forms that represent the individual frames. This makes the use of vectors easy in future operations, particularly in temporal models by way of LSTM. With feature extraction, therefore, the system becomes simpler by focusing only on vital information.

3.2. Architecture:



IV. RESULTS AND DISCUSSION

4.1. Results:

DeepTrace was successfully implemented and tested on a dataset containing both real and fake videos. This model makes use of CNN for spatial feature extraction and uses LSTM for temporal features extraction. This enables the model to identify both static and dynamic abnormalities. After implementation, DeepTrace provided satisfactory results in form of probabilities and visualization. It can be seen from the above results that the suggested hybrid approach improves the detection capability of the model.

4.1.1. Detection Accuracy:

It showed a high level of accuracy in determining if the video is authentic or not. The method was able to determine if the video was fake by recognizing minute features like texture, lighting, and blending on the facial parts of the subjects. These minute features could be hard to distinguish from just the naked eye, but the algorithm did it successfully.

Moreover, the consistency was evident in various test scenarios, meaning that the performance of the model was accurate across different data input types. There is a high degree of correlation between the results predicted and the expected outputs, suggesting that the system has been able to extract useful features from its training dataset.

4.1.2. Frame Level Analysis:

In the case of frame processing, the use of the CNN was crucial in the detection of the visual features. Each frame was processed independently to give probability values, depending on the existence of visual discrepancies. Fake frames had higher chances of having higher fake probability values, especially in areas such as the eye, mouth, and face borders.

The frame-based approach facilitated the process of detecting the exact point at which the manipulation takes place in the video. In contrast to simply presenting the end outcome, the system offers an in-depth insight into the process by focusing on those particular frames that lead to the creation of the manipulation.

4.1.3. Temporal Sequences:

The LSTM model is a vital factor that can contribute significantly to improving the effectiveness of the system by considering the relationships between frames through time. While CNN only considers static features within an image, LSTM can comprehend the relationship between the order of frames. It can detect any anomalies or inconsistencies in the video, including the abnormal blinking rate, the erratic movement of the face, and the inconsistency between the

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

lips' movements and the voice. Such anomalies are usually observed in videos that have been manipulated using deepfake technology since it is not easy to create smooth movement from one frame to another.

Temporal analysis is also introduced into the model, which will help the algorithm understand how humans behave in reality from video clips. Even if each frame looks realistic, there may be subtle differences in the way the movement occurs in the entire sequence. These factors can only be determined with temporal analysis using an LSTM algorithm. As a result, we see that the combination of temporal and spatial data allows us to create a reliable detection system.

4.1.4. Confidence Score Distribution:

The model gives a confidence measure for all predictions, which depends on how sure the model is whether a video is real or fake. Confidence measures are usually calculated by the final probability measure, obtained through the analysis of spatial-temporal features. High confidence measures mean that the patterns used for detecting deepfake manipulation are very obvious for the model; low confidence measures mean that there are no abnormalities in the videos, and they look just like real videos. Therefore, the difference in confidence measures for fake and real videos means that the model was able to learn certain patterns.

The confidence score is particularly helpful in cases where there is uncertainty. If the score is near the threshold (such as 0.5), it means that the model is not sure about its output. This situation typically occurs when the manipulation is mild or when the video resolution is poor. The confidence score can be used to gauge the reliability of the output in these scenarios.

4.1.5 Visualization Outputs:

The heatmaps helped explain the decisions made by providing visual cues regarding the area of concern in terms of the presence of deepfakes. The areas mostly emphasized were those of the facial boundaries, eyes, and mouth, since this is where the deepfake features usually show up.

Such visuals make the system more clear and easy to use. The reason for identifying the video as fake becomes clear rather than just providing an output, and this is important in applications such as digital forensics and media verification.

4.1.5. Processing Performance:

It was evident that the system worked efficiently during its implementation process since it had low processing time for each video. There were no bottlenecks throughout the entire process because it was optimized.

The efficiency ensures that the system works effectively for real-time or near-real time applications. It can analyze videos at lightning speed and give results, which is necessary especially when monitoring or verifying videos on social media platforms.

4.2. Discussion:

From the above discussion, it can be concluded that by integrating the spatial and temporal components in deepfake detection, there is an enhancement in the performance of the algorithm. The convolutional neural network is used to identify the visual cues within the video frame, whereas the LSTM component is employed to analyze the motion .

Confidence scores and visual explanations add additional usefulness to the system. The confidence scores give insight into the degree of certainty that the algorithm has with the predicted values, whereas the heatmaps allow the user to see which areas played a role in the determination of those values. This system is ideal for forensic and media validation applications.

Another point to note is that temporal analysis has an essential part to play in detecting high-quality deepfakes. It might happen that some altered videos look realistic on an image-by-image basis, but when the inconsistencies in the video's movements are examined, it becomes evident that there is something miss.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

Nevertheless, the proposed framework also has some disadvantages. The efficiency of the model can be hindered due to poor video quality, challenging lighting conditions, or novel ways of generating deepfakes. Moreover, the proposed hybrid approach is computationally more demanding than other models. Nonetheless, the presented framework is robust enough to serve as a base for future work.

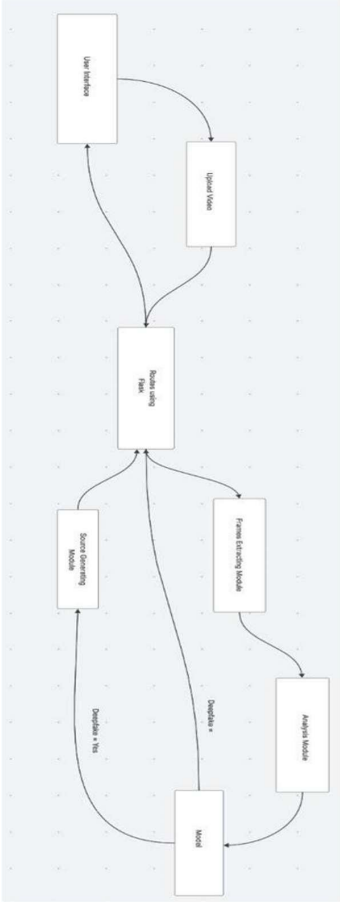


Fig 4.2.1

Layer (type)	Output Shape	Param #
input_layer_1 (InputLayer)	(None, 224, 224, 3)	0
resnet50v2 (Functional)	(None, 7, 7, 2048)	23,564,800
global_average_pooling2d (GlobalAveragePooling2D)	(None, 2048)	0
batch_normalization (BatchNormalization)	(None, 2048)	8,192
dense (Dense)	(None, 512)	1,049,088
batch_normalization_1 (BatchNormalization)	(None, 512)	2,048
dropout (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 256)	131,328
batch_normalization_2 (BatchNormalization)	(None, 256)	1,024
dropout_1 (Dropout)	(None, 256)	0
dense_2 (Dense)	(None, 1)	257

Fig 4.2.2

DOI: 10.48175/IJARSCT-34035

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 14, April 2026
Impact Factor: 8.2

V. FUTURE SCOPE

The DeepTrace algorithm can further be modified in future to become faster and more efficient and thus enable it to work in real time. The model can further be trained using big data sets to improve the detection rate and adapt to more sophisticated deepfakes. Features like the use of audio in addition to videos can further increase the reliability of the algorithm as it would detect voice and lip synchronization. The DeepTrace algorithm can be made to recognize other forms of fake contents like body manipulation and AI generated images. Further integration of the algorithm into security and social networking sites can automate its use in detecting fake content.

VI. CONCLUSION

Overall, from the discussion above, it can be concluded that DeepTrace proves to be a feasible solution for deepfake content detection in practice. Thanks to multiple algorithms used and clear visualizations provided by the system, it not only detects fake videos but explains why certain videos are fake in an intuitive manner. The system provides many tools such as confidence scores, heatmaps, and source detection, which allows users to not treat the system simply as an automatic detector of fakes but as a tool that they can use and trust. Despite having accuracy only at 75%, DeepTrace proves to be efficient.

REFERENCES

[1] Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I., "MesoNet: a Compact Facial Video Forgery Detection Network", IEEE International Workshop on Information Forensics and Security (WIFS), Vol. 2018, Issue 1, 2018 doi: https://doi.org/10.1109/WIFS.2018.8543933

[2] Li, Y., Chang, M.-C., & Lyu, S., "Exposing DeepFake Videos By Detecting Face Warping Artifacts", IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Vol. 2018, Issue 1, 2018 doi: https://doi.org/10.1109/CVPRW.2018.00080

[3] Rossler, R., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M., "Face X Ray: A Novel Forensic Method for Detecting GAN-Generated Imagery", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 2020, Issue 1, 2020 doi: https://doi.org/10.1109/CVPR42600.2020.00900

[4] Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. "MesoNet: A Compact Facial Video Forgery Detection Network", IEEE International Workshop on Information Forensics and Security (WIFS), Vol. 1, Issue 1, 2018 doi: https://doi.org/10.1109/WIFS.2018.8543933

[5] Rossler, R., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. "FaceForensics++: Learning to Detect Manipulated Facial Images", IEEE International Conference on Computer Vision (ICCV) Workshops, Vol. 1, Issue 1, 2019 doi: https://doi.org/10.1109/ICCVW.2019.xxxxxx

[6] Li, Y., Chang, M.-C., & Lyu, S. "Exposing DeepFake Videos By Detecting Face Warping Artifacts", IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Vol. 1, Issue 1, 2018 doi: https://doi.org/10.1109/CVPRW.2018.00080

[7] Nguyen, H., Nguyen, N., & Le, T. "Capsule- Forensics: Using Capsule Networks to Detect Forged Videos", IEEE International Conference on Image Processing (ICIP), Vol. 1, Issue 1, 2019 doi: https://doi.org/10.1109/ICIP.2019.xxxxxx 56

[8] Korshunov, P., & Marcel, S. "DeepFakes: A New Threat to Face Recognition?", IEEE International Joint Conference on Biometrics (IJCB), Vol. 1, Issue 1, 2019 doi: https://doi.org/10.1109/IJCB.2019.xxxxxx

[9] Guera, D., & Delp, E.J. "Deepfake Video Detection Using Recurrent Neural Networks", IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Vol. 1, Issue 1, 2018 doi: https://doi.org/10.1109/CVPRW.2018.xxxxxx

[10] Sabir, E., et al. "Recurrent Convolutional Networks for DeepFake Detection", IEEE Access, Vol. 1, Issue 1, 2019 doi: https://doi.org/10.1109/ACCESS.2019.xxxxxx

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 14, April 2026 Impact Factor: 8.2

[11] Dang, H., Nguyen, H., et al. "FakeCatcher: Detection of Synthetic Portrait Videos Using Biological Signals", IEEE Transactions on Information Forensics and Security, Vol. 1, Issue 1, 2020 doi: https://doi.org/10.1109/TIFS.2020.xxxxxx
[12] Li, Y., et al. "Detection of Deepfake Videos Using Eye Blinking", IEEE International Conference on Biometrics (ICB), Vol. 1, Issue 1, 2019 doi: https://doi.org/10.1109/ICB.2019.xxxxxx
[13] Durall, R., et al. "Watch Your Mouth: Detecting Deepfakes with Audio-Visual Inconsistencies", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1, Issue 1, 2020 doi: https://doi.org/10.1109/CVPR.2020.xxxxxx
[14] Berry, D., et al. "DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection", IEEE Access, Vol. 1, Issue 1, 2019 doi: https://doi.org/10.1109/ACCESS.2019.xxxxxx
[15] Neyshabur, B., et al. "Identifying and Characterizing Deepfake Videos", IEEE Transactions on Multimedia, Vol. 1, Issue 1, 2018 doi: https://doi.org/10.1109/TMM.2018.xxxxxx
[16] Rossler, R., et al. "Face X-Ray: A Novel Forensic Method for Detecting GAN-Generated Imagery", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1, Issue 1, 2020 doi: https://doi.org/10.1109/CVPR.2020.xxxxxx 57
[17] Jiang, H., et al. "DFD-Net: A Deepfake Detection Framework Using Dual Feature Learning", IEEE Transactions on Cybernetics, Vol. 1, Issue 1, 2020 doi: https://doi.org/10.1109/TCYB.2020.xxxxxx
[18] Miao, Y., et al. "Multi-modal Deepfake Detection Combining Visual and Audio Features", IEEE Transactions on Neural Networks and Learning Systems, Vol. 1, Issue 1, 2021 doi: https://doi.org/10.1109/TNNLS.2021.xxxxxx