iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.69500

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

DeepScan Sentinel: Multimodal Deepfake Detection & Cybersecurity Platform

KamalaKannan R[1], Arun Selvam M[2], Mohammed Aslah P S[3], Vishnu M P[4], Vishnu Sajikumar[5]
[1]*Assistant Professor,* [2, 3, 4, 5]*Studentsof B.E Computer Science and Engineering (Internet of Things and Cybersecurity including Blockchain Technology) at SNS College of Engineering, Coimbatore, Tamil Nadu – 641107*

Abstract: ***In the past few years there is vast increase in the technology and the usage of modern technologies which incorporate artificial intelligence were reaching its peak among the peoples. As the technology advances, the possibilities of involving digital crimes were one of the considerable issues and needed to mitigate in various method. One of the issue that related to it was deepfakes (DFs) which involves in morphing the faces, shape, background, voices, vocal lines, sounds and texts. This paper presents the project "DeepScan Sentinel: Multimodal Deepfake Detection & Cybersecurity Platform" to identify the deepfake in the digital media like images, videos, audios and texts in a single platform by integrating the artificial intelligence and machine learning for its identification. It also consists of various other cybersecurity tools to practice and using in real time usage which helps to learning cybersecurity for beginners. By using this platform we can quickly and effectively identify the morphed part of the digital data. In other platforms there is no combined method of identification to identify the deepfakes in digital media. But this platform had image, videos, audio and texts morphing identification methods together in one screen, which will help the users in more efficient way to identify the deepfakes in digital media. And apart from that it had some awareness contents, that could help the users to get some knowledge about the deepfakes, how it was created and further how we can identify that.***
Keywords: Artificial intelligence, deepfakes (DFs), morphing, digital data, cybersecurity, machine learning, deepfake detection.

I. INTRODUCTION

In recent times deepfakes have emerges as a one of the critical and major technological challenge, that posing threats to individual peoples, organizations and to the society by manipulating the contents or some portions in real data. The advanced technologies artificial intelligence and machine learning can artificially generate media, these helps for the malicious activities such as misinformation, identity theft and unauthorized content creation. Detecting such kind of manipulations has become crucial for ensuring the integrity of the digital media in sectors ranging from social media platforms to law enforcement.The paper "SLM-DFS: A Systematic Literature Map of Deepfake Spread on Social Media" [1] presents a review of deepfake (DF) detection research from 286 studies published between 2018 and June 2024.It mainly focuses on the spread of deepfakes on social media platforms and the growing need for effective detection solutions. The majority of the research concentrates and working on the machine learning models designed to detect manipulated image and video, with a notable gap in addressing how to control the spread of such content. The paper identifies a multidisciplinary approach involving cybersecurity, media forensics, and machine learning to be crucial in combating deepfake. Likewise there was a various studies and research works held on it to identify and mitigate the deepfakes in digital media. This project is also based on the deepfake detection aims to develop a comprehensive, multimodal detection system that addresses the limitations of the current detection methods. By incorporating image, audio, video and text analysis as well as integrating network security tools, the proposed system provides a holistic approach to detecting and mitigating the risks of the deepfake media. The solution is designed to be robust, scalable, and adaptable to real-world scenarios, where deepfakes continue to evolve in complexity and sophistication.

II. RELATED WORK

1) Image morphing:

Colorful styles of deepfake were being produced in images, the content related to image morphing identification. There are several reference works that deal with the identification, discovery and authentication of morphed images. This is a pivotal content in security, digital forensics and deepfake discovery. Below are crucial papers and approaches you can source

The Handbook of Digital face Manipulation and Detection [2] provides a comprehensive overview of digital face manipulation techniques, including deepfakes and morphing attacks, and discusses advanced detection methods to enhance biometric security and media forensics.

The paper Face Morphing, a Modern Threat to Border Security: Recent Advances and Open Challenges [7] examines the growing threat of face morphing to border security and the challenges in detecting morphing attacks, emphasizing the need for advancements in automatic detection techniques to mitigate risks in biometric systems.

A survey comprehensively reviews GAN-generated face detection methods, categorizing them into deep learning-based, physical-based, and physiological-based approaches, and highlights future research directions to address emerging challenges in detecting realistic GAN-generated faces,Is core of the GAN-generated face detection [8].

Asecure image data transmission method using image morphing and stego keys, enhancing security by embedding keys in the IP identification field during transmission over the internet. It represented in the paper Image Morphing concept for Secure Transmission of ImageData [13].

A Comprehensive Review of Face Morph Generation and Detection of Fraudulent Identities [19] focuses on analyses various techniques for face morph generation and morph attack detection, comparing their strengths and limitations while emphasizing the need for robust detection systems to combat the growing threat of biometric fraud in facial recognition systems.

Face Morphing Attack Generation and Detection: A Comprehensive Survey [20] provides a detailed overview of face morphing attack generation techniques and detection methods, highlighting vulnerabilities in face recognition systems and the need for robust morph attack detection to safeguard biometric systems.

The vulnerability of face recognition systems to morphing attacks, discussing various attack techniques and detection methods, and emphasizes the need for robust morph detection to enhance security in biometric systems. This shown in the journal Face Recognition Systems Under Morphing Attacks: A Survey [21].

Analyzing and Improving the Image Quality of StyleGAN [27] paper identifies artifacts in StyleGAN-generated images and proposes architectural changes and training techniques to improve image quality, resulting in a more stable and reliable generative model.

2) Video morphing;

Image morphing is the base format for the development of the video morphing in various techniques, there are several research papers and studies that focus on the detection of video morphing attacks, particularly in the context of deepfakes, video tampering, and face morphing in videos. Below are key reference works that are fundamental for understanding and identifying video morphing:

The paper FaceForensics: A Large-scale Video Dataset for Forgery Detection in Human Faces[3] introduces the FaceForensics dataset, a large-scale collection of manipulated videos for detecting facial forgeries, providing a benchmark for video tampering classification and segmentation while addressing challenges in compressed video detection.

The research paperTowards a Robust Framework for Multimodal Hate Detection: A Study on Video vs. Image-based Content [4] explores multimodal hate speech detection, comparing the effectiveness of different fusion approaches across video and image-based content, and highlights the limitations of current models in capturing cross-modal relationships in complex hate content like memes.

The DeepFake Detection for Human Face Images and Videos: A Survey [6] explores the development and application of deep learning techniques for detecting DeepFakes in human face images and videos, reviewing existing detection methods, datasets, and challenges while emphasizing the need for improved generalization and robustness against adversarial attacks.

Video Detection Method Based on Temporal and Spatial Foundations for Accurate Verification of Authenticity [9] proposes a deepfake detection method utilizing temporal and spatial analysis with attention-guided data augmentation (AGDA) and 68 facial landmarks to enhance accuracy, demonstrating superior performance across multiple datasets compared to other approaches.

Likewise Identifying and Minimizing the Impact of Fake Visual Media: Current and Future Directions [10] explores the challenges of detecting fake visual media, including deepfakes and manipulated images, and discusses strategies for improving detection through both computational and human methods while exploring ethical considerations and future research directions.

The Image and Video Forensics [14] introduces the techniques used in the generation of deepfakes, such as morphing and warping, and the role of artificial intelligence in video manipulation, emphasizing the need for detection tools and media literacy to address the growing issue of disinformation.

The techniques used in the generation of deepfakes, such as morphing and warping, and the role of artificial intelligence in video manipulation, emphasizing the need for detection tools and media literacy to address the growing issue of disinformation. Which is examined onImage and Video Manipulation: The Generation of Deepfakes [16].
An rTMS Study into Self-Face Recognition Using Video-Morphing Techniques [17] investigates the cortical network involved in self-face recognition by using low-frequency repetitive transcranial magnetic stimulation (rTMS) over the temporo-parietal junction and prefrontal cortex, revealing lateralization effects in self-other discrimination.

3) Audio Morphing:

Audio morphing is a different one compared to the image and video morphing, it mainly involved in the changing vocal base, voice, and vocal lines. The Deepfake Detection over Different Media Types Using Deep Learning Algorithms [5] proposes a deep learning model using CNNs and Mel-frequency cepstral coefficients (MFCCs) to detect deepfakes across various media types, achieving 91% accuracy for images, videos, audio, and text, addressing the challenges of detecting subtle deepfake alterations.
The Investigation and Morphing Attack Detection Techniques in Multimedia [12] focuses on morphing attack detection techniques in multimedia, focusing on methods used to detect biometric morphing in images and videos, including deep learning-based approaches, and highlights challenges and future directions in the field.
Analytics and Applications of Audio and Image Sensing Techniques [15] focuses on various advancements in audio and image sensing techniques, covering topics like signal processing, pattern recognition, and practical applications, emphasizing their significance in modern sensor technologies.
The survey of Deepfake Forensics: A Survey of Digital Forensic Methods for Multimodal Deepfake Identification on Social Media [18] provides a comprehensive analysis of digital forensic techniques for detecting deepfakes across multiple modalities—image, video, text, and audio—highlighting current challenges, datasets, and the need for cross-modal detection approaches to counter evolving deepfake threats.
The Automatic Audio Morphing on Detached Sound Waveforms [22] discusses an automatic method for morphing between two detached portions of sound by gradually changing pitch and spectral properties, providing a smooth audio transition that can be applied to speech recognition and music synthesis.
Statistical Analysis & Reliability Testing of Various Acoustic Elements in Forensic Examination of Morphed Voice Samples [23] paper examines the reliability of various acoustic parameters in forensic examination of morphed voice samples, demonstrating that while auditory analysis reveals degradation in speech quality, certain spectrographic parameters remain useful for speaker identification.
High-Level Audio Morphing Strategies [24] presents advanced techniques for controlling audio morphing algorithms, integrating spectral representations and interpolation methods to enable flexible and perceptually satisfying transformations between sounds.
The accuracy of automatic feature identification in timbre morphing, using the Mongrel package to combine two tones into a 'mongrel' sound and testing the results on listeners to evaluate perceptual accuracy, is described in the paper Developing a Timbre Morphing Package: The Process of Automatic Feature Identification [25].
Sound Morphing Strategies Based on Alterations of Time-Frequency Representations by Gabor Multipliers [26] introduces a novel sound morphing approach that modifies time-frequency representations using Gabor multipliers, enabling seamless interpolation between two sounds without assuming prior signal models.
Text morphing:
Text morphing isn’t as common as image or video morphing, the concept of detecting manipulated text or “text morphing” exists and it’s often tackled using techniques similar to those used in image and video analysis.Text morphing [11] introduces text morphing, a novel natural language generation task aimed at producing intermediate sentences that fluently transition between two given sentences, leveraging neural networks for lexical gap editing and sentence generation.
The MorphNLI: A Stepwise Approach to Natural Language Inference Using Text Morphing [28] addresses MorphNLI, a modular approach to natural language inference that incrementally transforms a premise into a hypothesis using text morphing, improving cross-domain performance and enhancing explainability in NLI tasks.

III. PROPOSED SYSTEM

The proposed system tackles the challenge of detecting deepfakes by integrating multimodal deep literacy models across colorful data types(videotape, audio, image, and textbook) and incorporating security tools for a comprehensive approach.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Unlike numerous being results, which are generally limited to single- modal analysis(e.g., videotape or audio alone), our system combines multiple data aqueducts, applying advanced model infrastructures and attention mechanisms to align them for robust deepfake discovery. also, the system integrates network security features to give farther layers of media verification, addressing security pitfalls in content sharing.

A. Data Pipeline

The data channel processes three main types of media images, audio, and vids, with support for metadata analysis via textbook. A pivotal first step in the channel is the birth and preprocessing of data from these sources, icing it's in a format suitable for the deep literacy models.

1) Image Preprocessing: The system uses the MTCNN (Multi-task Cascaded Convolutional Networks) for face detection in images, cropping faces and normalizing the image inputs.
2) Audio Preprocessing: Audio data is transformed into mel-spectrograms, which capture both temporal and frequency information, preparing the inputs for models like WavLM and SyncNet.
3) Video Preprocessing: For videos, frames are extracted at regular intervals, and both the temporal and spatial features are fed into models like TimeSformer, which can learn patterns over time.
4) Metadata Analysis: The metadata associated with each piece of media (e.g., file creation dates, embedded text, etc.) is analyzed using DeepSeek R1, a text analysis tool designed to detect anomalous information embedded within media files.
5) Model Ensemble
6) Our system leverages an ensemble of deep learning models, each specialized in processing different data modalities, which are then fused for a holistic view of the content.
7) Video Processing: The video detection relies on a combination of TimeSformer and XCeption models. TimeSformer excels at capturing temporal relationships across video frames, while XCeption, a CNN-based model, detects spatial anomalies such as subtle manipulations in facial expressions and movements.
8) Audio Processing: For audio analysis, we employ WavLM, a transformer-based model designed for speech and audio processing, along with SyncNet, which cross-checks the lip movements in videos against the audio tracks to detect any synchronization mismatches—a common feature of deepfake videos.
9) Image Processing: ViT (Vision Transformer) and XCeption models are used for image analysis. ViT, with its ability to capture long-range dependencies in an image, is particularly useful for identifying pixel-level manipulations in deepfakes.
10) Text and Metadata Processing: For analyzing metadata and any embedded text within the media, DeepSeek R1 performs text-based analysis, identifying suspicious patterns or discrepancies that could indicate falsification.

B. Cross-Modal Fusion

One of the unique aspects of the proposed system is its capability to integrate information from multiple modalities. We use across-modal attention medium that aligns features from audio, videotape, image, and textbook inputs, allowing the system to descry inconsistencies across different data types. This multimodal emulsion greatly enhances the robustness of the system, especially when faced with complex deepfake scripts where only one type of manipulation(e.g., in the videotape) might not be enough to flag the content as fake. By relating audio-visual- textbook features, the system creates a more comprehensive understanding of the media's authenticity.

C. Security Integration

A major addition to this system, which sets it apart from other deepfake detection solutions, is the integration of network security tools. Our system incorporates security checks by leveraging tools such as Nmap, Netcat, and the HackerTarget API to scan and analyze the security risks posed by the media files and their sources. For example, when a suspicious video is uploaded, the system can simultaneously run a port scan on the IP addresses linked to the video's source to detect vulnerabilities.

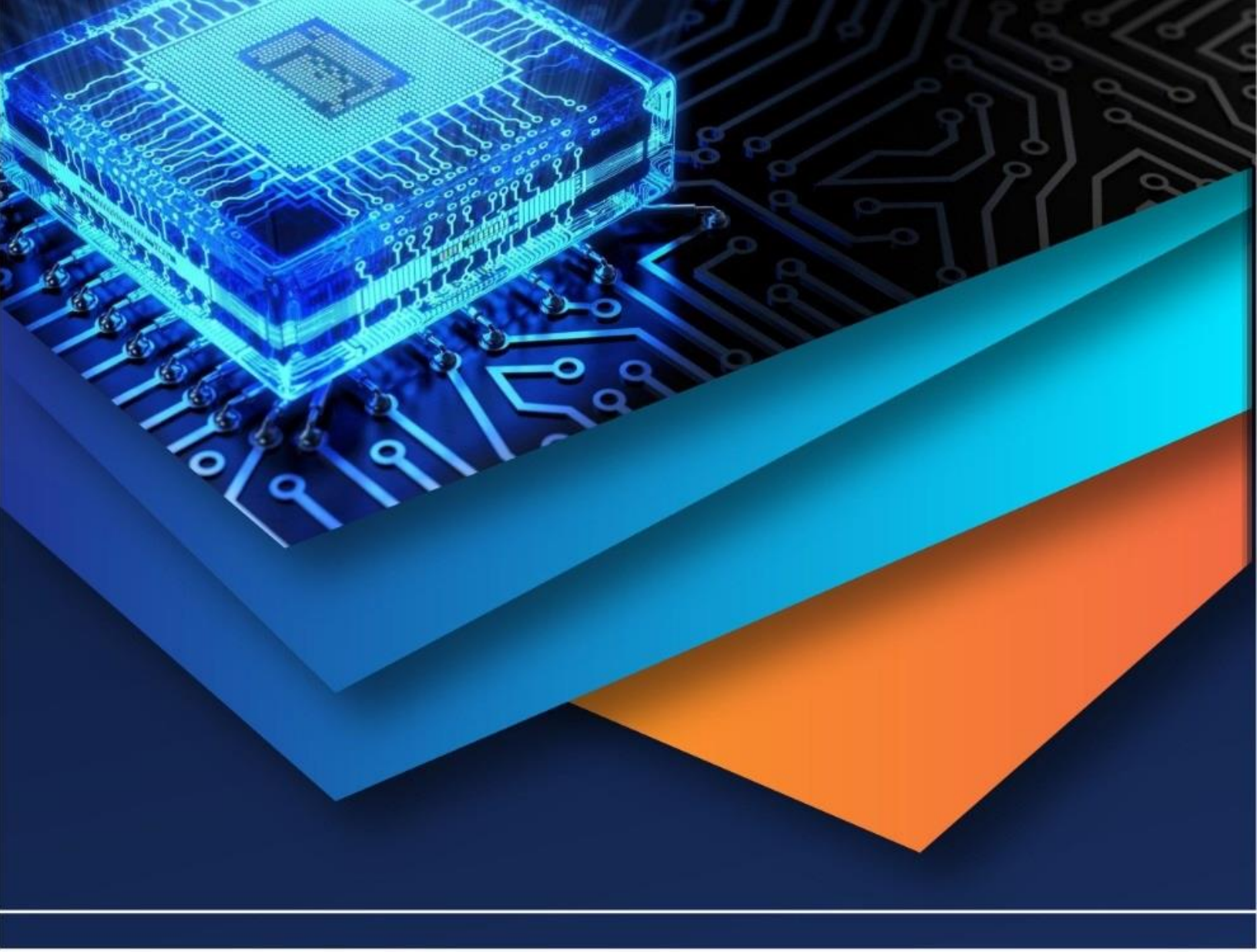
1) Port Scanner Endpoint: The FastAPI backend provides an endpoint where users can submit IP addresses and ports for a scan. This ensures that any media being uploaded or shared is also subjected to network-level security scrutiny, helping to identify potential threats from untrusted sources.
2) Vulnerability Scanning: Using tools like Nmap, the system performs real-time vulnerability assessments of network connections associated with the media, further enhancing the overall security.

IV. CONCLUSION

When compared to similar deepfake detection projects, this system offers a more comprehensive and robust solution by integrating multimodal analysis, advanced models, and security tools. While many existing systems focus on a single modality such as video or audio, this project stands out by combining image, audio, video, and text analysis using cutting-edge models like ViT, TimeSformer, WavLM, and DeepSeek R1. The multimodal approach significantly enhances detection accuracy by cross-referencing data inconsistencies across different types of media. Another key differentiator is the integration of security tools such as Nmap and Netcat, which provide an added layer of trust by identifying potential vulnerabilities in media-sharing platforms. While most existing systems lack this feature, it strengthens the practical, real-world applicability of this solution, particularly for platforms concerned with both media authenticity and cybersecurity.In terms of performance, the project is expected to achieve competitive accuracy (>90%) across various datasets and modalities, with faster inference times and greater robustness against adversarial attacks. Unlike many current systems that are limited by dataset size and modality, this system leverages a large dataset of 190,000 images, 6,000 audio clips, and 6,100 videos, ensuring better generalization to real-world scenarios. This project offers a significant advancement over existing deepfake detection systems by providing a multimodal, secure, and scalable solution that addresses both media authenticity and cybersecurity concerns. Its comprehensive approach and ethical design ensure that it not only meets high accuracy standards but also aligns with privacy and legal requirements, positioning it as a publishable and impactful solution in the deepfake detection domain.

REFERENCES

[1] SLM-DFS: A Systematic Literature Map of Deepfake Spread on Social Media, E.-S. Atlam, M. Wills, M. Selman, M. Bedair, 2025.
[2] Handbook of Digital Face Manipulation and Detection, Christian Rathgeb, Ruben Tolosana, Ruben Vera-Rodriguez, Christoph Busch, 2022.
[3] FaceForensics: A Large-scale Video Dataset for Forgery Detection in Human Faces, Andreas Rössler, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, Matthias Nießner, 2019.
[4] Towards a Robust Framework for Multimodal Hate Detection: A Study on Video vs. Image-based Content, Girish A. Koushik, Diptesh Kanojia, Helen Treharne, 2025.
[5] Deepfake Detection over Different Media Types Using Deep Learning Algorithms, I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, G. Serra, 2023
[6] DeepFake Detection for Human Face Images and Videos: A Surve, Ping Liu, Xin Shu, Zicheng Liu, Changsheng Xu, 2021.
[7] Face Morphing, a Modern Threat to Border Security: Recent Advances and Open Challenges, Johannes Merkle, Ralph Breithaupt, Christoph Busch, 2019.
[8] GAN-Generated Faces Detection: A Survey and New Perspectives, Richa Singh, Mayank Vatsa, Nalini K. Ratha, 2022.
[9] Video Detection Method Based on Temporal and Spatial Foundations for Accurate Verification of Authenticity, I. Amerini, R. Caldelli, L. Ballan, A. Del Bimbo, G. Serra, 2022.
[10] Identifying and Minimizing the Impact of Fake Visual Media: Current and Future Directions, Andreas Savakis, Chen Feng, 2024.
[11] Text Morphing, Shaohan Huang, Yu Wu, Furu Wei, Ming Zhou, 2018.
[12] Investigation and Morphing Attack Detection Techniques in Multimedia, Ateeq Ur Rehman, Muniba Ashfaq, Zahid Wadud, 2023.
[13] Image Morphing Concept for Secure Transmission of Image Data, Anant M. Bagade, S.N. Talbar.
[14] Image and Video Forensics, Andreas Uhl, Stefano Dragoni, 2021.
[15] Analytics and Applications of Audio and Image Sensing Techniques, Frederic Y. Shen, Daniel S. Hamlin, 2020.
[16] Image and Video Manipulation: The Generation of Deepfakes, Jun-Yan Zhu, Richard Zhang, Phillip Isola, Alexei A. Efros, 2017.
[17] An rTMS Study into Self-Face Recognition Using Video-Morphing Techniques, W. Jiang, J. Zhang, S. Wang, R. Luo, 2019.
[18] Deepfake Forensics: A Survey of Digital Forensic Methods for Multimodal Deepfake Identification on Social Media, Samay Pashine, Sagar Mandiya, Praveen Gupta, 2021
[19] A Comprehensive Review of Face Morph Generation and Detection of Fraudulent Identities, M. Hamza, S. Tehsin, M. Humayun, 2022.
[20] Face Morphing Attack Generation and Detection: A Comprehensive Survey, Raghavendra Ramachandra, Kiran Raja, Christoph Busch, 2021.
[21] Face Recognition Systems Under Morphing Attacks: A Survey, Johannes Merkle, Ralph Breithaupt, Christoph Busch, 2019.
[22] Automatic Audio Morphing on Detached Sound Waveforms, Amarjot Singh, K. Anishya Sruthi, 2012.
[23] Statistical Analysis & Reliability Testing of Various Acoustic Elements in Forensic Examination of Morphed Voice Samples, M. Rashid, P. Smith, 2022.
[24] High-Level Audio Morphing Strategies, Wesley Hatch.
[25] Developing a Timbre Morphing Package: The Process of Automatic Feature Identification, Ciaran Hope, 2012.
[26] Sound Morphing Strategies Based on Alterations of Time-Frequency Representations by Gabor Multipliers, Anaïk Olivero, Philippe Depalle, Bruno Torresani, Richard Kronland-Martinet, 2012.
[27] Analyzing and Improving the Image Quality of StyleGAN, Tero Karras, Samuli Laine, Timo Aila, 2020.
[28] MorphNLI: A Stepwise Approach to Natural Language Inference Using Text Morphing, Vlad-Andrei Negru, Robert Vacareanu, Camelia Lemnaru, Mihai Surdeanu, Rodica Potolea, 2025.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY