International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

Advanced Deep Fake Detection System

Kumpatla Chinna[1], Bolla Siva Naga kumar [2], Somisetti Satya[3], Polisetti S.S.C. Mahesh[4], Bokka Satish[5], Alapati Ramakrishna[6]

[1,2,3,4,5,6]Department of Artificial Intelligence and Machine Learning, Bonam Venkata Chalamayya Engineering College, Andhra Pradesh, India

Email:

Abstract

The production and distribution of deepfake material in text, audio, video, and image modalities has greatly grown because to the quick development of generative artificial intelligence. Public safety, journalism, cybersecurity, and digital trust are all seriously threatened by such synthetic media. The efficiency of current detection techniques in handling intricate, multimodal manipulation techniques is limited by their primary focus on single-modality analysis. This work offers an integrated multimodal deepfake detection methodology intended to evaluate authenticity across diverse media sources inside a single architecture in order to get over this restriction. A Convolutional Neural Network (CNN) is used to extract spatial information from video frames and detect manipulation traces, visual artifacts, and face abnormalities in order to detect video deepfakes. In order to detect audio deepfakes, discriminative acoustic features are extracted and then classified using a Random Forest method, which offers resilience against attacks using speech synthesis and voice cloning. A TF-IDF Vectorizer in conjunction with Cosine Similarity is used to quantify the semantic similarity between reference materials and input text in order to detect textual plagiarism. The s ystem incorporates a refined vision transformer model based on SIGLIP for image authenticity verification in order to differentiate between AI-generated images and content provided by humans. Furthermore, a RoBERTa-base transformer model that can categorize both machine-generated and human-written text utilizing contextual embeddings is used for AI-generated text identification. A full authenticity score is generated by combining the outputs from all modalities. The framework's usefulness for automated content moderation systems and real-world digital forensics is highlighted by experimental evaluation, which shows better robustness and scalability in comparison to isolated detection approaches.

Keywords

Deepfake Detection, Multimodal Learning, Convolutional Neural Network (CNN), Random Forest Classifier, TF-IDF Vectorizer, Cosine Similarity, Vision Transformer, SIGLIP, RoBERTa, AI-Generated Content Detection, Digital Forensics, Content Authentication, Machine Learning

1. Introduction

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

The creation and consumption of digital content have changed dramatically as a result of artificial intelligence's explosive rise, especially in generative modeling. These days, sophisticated deep learning models may produce incredibly lifelike text, audio, video, and image content that nearly resembles real human-produced data. Although these technologies have a lot to offer in terms of automation, education, and entertainment, they also present substantial hurdles in the form of deepfakes, which are artif icially created

or altered material that are meant to look real. Critical issues about disinformation, identity theft, financial fraud, political manipulation, and the decline in public trust in digital platforms have been brought up by the misuse of such content.

Thus, deepfake detection has emerged as a key field of study in cybersecurity and digital forensics. Conventional detection techniques, such text-based AI content categorization or video-based facial artifact detection, usually concentrate on a single data modality. However, contemporary deepfake attacks are progressively combining various modalities, such as AI-generated captions or transcripts and synthetic speech timed with altered footage. In such complicated situations, single-modality systems frequently fall short of providing accurate detection, underscoring the necessity of an all-encompassing and integrated strategy.

This study suggests a multimodal deepfake detection framework that can analyze text, audio, video, and images in a single sys tem in order to overcome this difficulty. To guarantee reliable identification across a variety of media sources, the system makes use of transformer-based structures, deep learning models, and traditional machine learning methods. The suggested system improves scalability, flexibility, and dependability to changing generative AI technologies by combining predictions from several anal ytical pipelines. The goal of this effort is to aid in the creation of workable and efficient solutions for practical content authentication and digital security applications.

Fig-1 Multimodal Deepfake Detection Framework

The suggested system is a multimodal deepfake detection framework intended to authenticate digital information in the text, image, audio, and video domains, as seen in Fig. 1. User-uploaded media can be processed by the framework using certain analytical modules. Specialized algorithms are used by each module to identify manipulation artifacts and extract significant features. To guarantee precise and dependable identification, the system combines transformer-based models, deep learning, and traditional machine learning.

A. Rossler, et.al,[1] The realism of contemporary facial alteration techniques and the difficulties in both automatic and hum an detection are examined in this research. In order to assess deepfake detection techniques, it presents a sizable benchmark dataset with more than 1.8 million altered photos and shows that data-driven detectors with domain expertise can perform better than human observers. D. Afchar, et.al,[2] In order to identify face tampering in films caused by Deepfake and Face2Face approaches, this study suggests a deep learning method. The technique achieves excellent detection accuracy of over 98% for Deepfake and 95% for Face2Face using lightweight neural networks. Y. Li, et.al,[3] M2U-Net, an effective neural network design for retinal vascular segmentation that outperforms current techniques, is proposed in this research. It enables real -time processing and deployment on mobile and embedded systems by utilizing a lightweight encoder–decoder design with MobileNetV2 components.

Y. Li, et.al,[4] This study suggests a technique for identifying deepfake videos by examining eye blinking patterns, a physiological

indicator that is frequently absent from artificially produced fake videos. The method successfully detects videos made with deep neural network-based DeepFake techniques by evaluating eye-blink detection. D. Güera, et.al,[5] This study presents a temporal - aware deepfake detection system that analyzes temporal patterns in films using an RNN and extracts frame-level features using a CNN. Using a straightforward architecture, the method successfully detects faked films and produces competitive results on deepfake datasets. C. N. Duong, et.al,[6] MobiFace, a lightweight deep neural network with low memory and computational needs, is proposed in this paper for effective facial recognition on mobile devices. In comparison to big deep networks, the model greatly reduces processing time and memory use while achieving high accuracy of 99.73% on LFW and 91.3% on MegaFace. J. Thies, et.al,[7] A real-time facial replication technique that copies facial expressions from a source video to a target video is presented in this study. Photo-realistic modified movies that seamlessly integrate with the original scenario are produced by the system. Y. Nirkin, et.al,[8] This study presents Face switching GAN (FSGAN), a deep learning technique that does reenactment and face switching without the need for face-specific training. It creates realistic switched faces while maintaining lighting and skin ton e using RNN-based reenactment, face completion, and blending networks. K. P. Nandan, et.al,[9] Using models like LSTM, ALBERT, FNNet, and CNN+RNN, this article suggests a deep learning-based solution for real-time fake news identification. With ALBERT obtaining 100% and CNN+RNN achieving 99%, the results demonstrate great accuracy, demonstrating their efficacy in spotting fake news articles. V. C. Handaragall, et.al,[10] This survey uses AI techniques like machine learning and deep learning to examine several deepfake image and video detection

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

methods. It examines methods such as DST-Net, DMHLP, and VARMA-LSTM-GRU, emphasizing their effectiveness and difficulties in precisely identifying altered media.

2. Method

A multimodal deepfake detection framework that can analyze many digital media types, such as photos, videos, audio, text, and documents, is part of the suggested system. Many machine learning and deep learning models, each tailored for a certain data modality, are included into the system. The pipeline starts with media input and preprocessing, moves on to modality - specific analysis using trained models, and ultimately provides an authenticity classification that indicates whether the information is authentic or fake. This is the suggested approach shown in Figure 2.

User Authentication and Security Module

To guarantee safe access to the deepfake detection platform, the suggested solution includes a user authentication mechanism. A login and signup module is used to implement authentication, confirming user identification prior to granting access to system features. Only registered users are able to use the detection services thanks to this method, which also helps prevent unwanted access to the system. New users must enter their username, email address, and password as part of the registration procedure. The user data is safely stored in the database after the system verifies the input data. Before being stored, passwords are encry pted using secure hashing algorithms to improve security and prevent critical user credentials from being directly exposed.

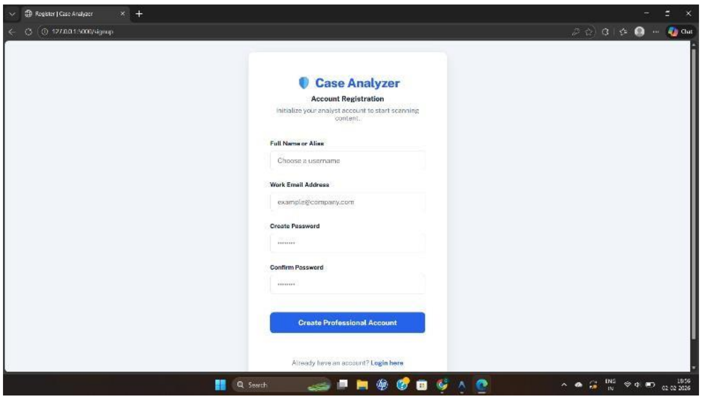


Fig 2: User Registration UI

Users must input their registered credentials during the login procedure, which are subsequently checked against the database entries that have been stored. Access is only granted if the credentials match the encrypted data that has been stored after the authentication module verifies the legitimacy of the supplied username and password. The system rejects access and asks the us er to reenter correct credentials if they are entered incorrectly.

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

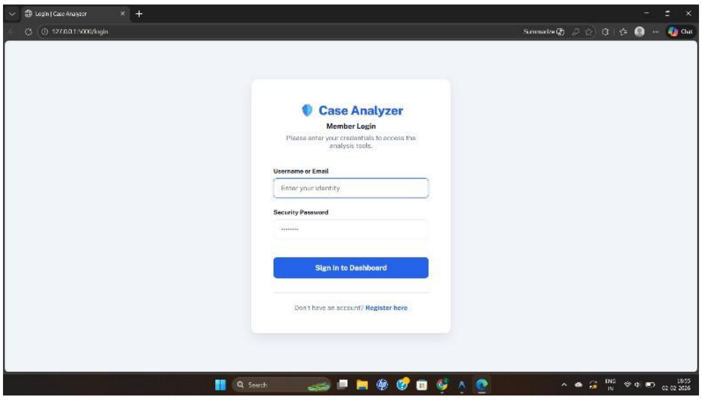


Fig 3: User Login UI

This authentication layer guards against unauthorized use of the system and guarantees data privacy. The system preserves regulated access to the deepfake detection framework while protecting user data and upholding system integrity by incorporati ng a secure login and signup process.

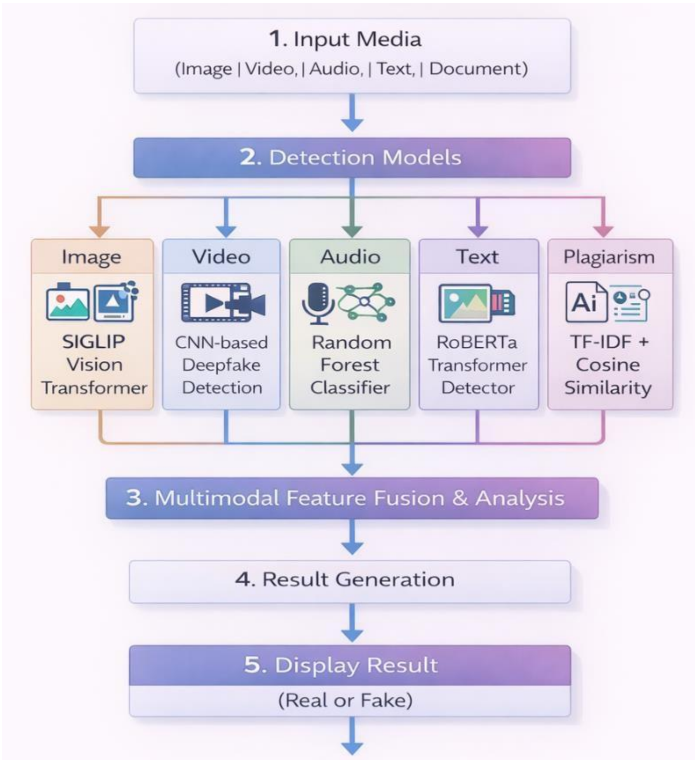


Fig 4: Proposed Method

2.1 Input Media

The suggested multimodal deepfake detection method starts with the Input Media stage. In this

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

phase, the system uses a web interface to receive different types of digital content from the user.

This stage's main objective is to gather and classify the uploaded media so that the appropriate detection algorithm can process it.

For plagiarism analysis, the system supports a wide range of input data forms, such as images, videos, audio files, text, and documents. The system can evaluate a wide range of digital content, including stuff that has been updated or created by artificial intelligence, thanks to its multimodal functionality. Support for diverse input formats significantly increases detection eff icacy because deepfake technology may alter a wide variety of media kinds.

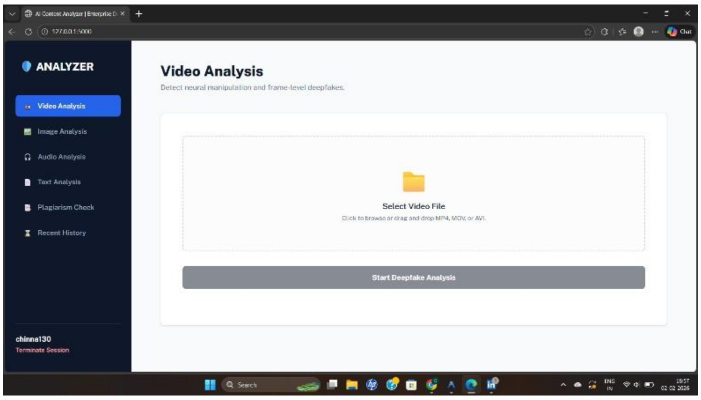


Fig 5: Deep Fake System Interface

The system uses the file format to detect the type of media when a user uploads a file. For instance, video files can be in M P4 or AVI formats, while image files can be in JPG, PNG, or JPEG formats. While text inputs might be simple text or document files, audio inputs are typically permitted in WAV or MP3 formats.

This step is implemented utilizing Python-based backend technologies, as well as web frameworks like Flask or FastAPI to handle HTTP requests and file uploads. The uploaded files are momentarily stored on the server and then routed to the appropriate processing pipeline. Each pipeline is particularly intended to examine a certain media type, employing specialized machine learnin g or deep learning models.

For example, image inputs are routed to the Vision Transformer-based SIGLIP model, video inputs are processed by a Convolutional Neural Network (CNN) model, audio files are analyzed by a Random Forest classifier, and text inputs are evaluated by a RoBERT a transformer model. Additionally, documents are examined for plagiarism using TF-IDF vectorization and cosine similarity algorithms.

Overall, the Input Media stage serves as the system's entrance point, allowing for the flexible and efficient handling of various types of digital content before moving on to the preprocessing and analysis stages of the proposed deepfake detection framework.

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

2.2 Models

2.2.1 Image Detection using SIGLIP Vision Transformer

The image analysis module utilizes the SIGLIP Vision Transformer, a transformer-based architecture designed for robust visual representation learning:

1. The input image is first preprocessed by performing resizing and normalization to standardize the image and improve the performance of the detection model.
2. The preprocessed image is divided into fixed-size patches so that the model can analyze different regions of the image separately.
3. Each image patch is converted into a numerical vector representation known as an embedding, which allows the transformer model to process the visual information.
4. These embeddings are then passed through multiple transformer encoder layers that consist of multi-head self-attention mechanisms and feed-forward neural networks.
5. The self-attention mechanism analyzes relationships between different image patches and helps the model identify spatial inconsistencies and abnormal patterns commonly found in deepfake images, such as unnatural facial textures, irregular lighting, and distorted facial structures.
6. The SIGLIP model further improves detection by learning semantic relationships between visual features and contextual information, enabling the identification of subtle manipulation artifacts that may not be detected by traditional CNN-based models.
7. Finally, the extracted features are passed through a classification layer that determines whether the image is authentic or manipulated (deepfake).

2.2.2. Video Deepfake Detection using CNN

The video deepfake detection module uses a Convolutional Neural Network (CNN) to examine the visual details in video frames. Deepfake videos often create small visual errors when faces are swapped or generated. CNN models are good at finding these errors because they can automatically learn important visual features like edges, textures, and facial patterns. The model checks each frame of the video and analyzes these features. By combining the results from multiple frames, the system can decide whether the video is real or fake.

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

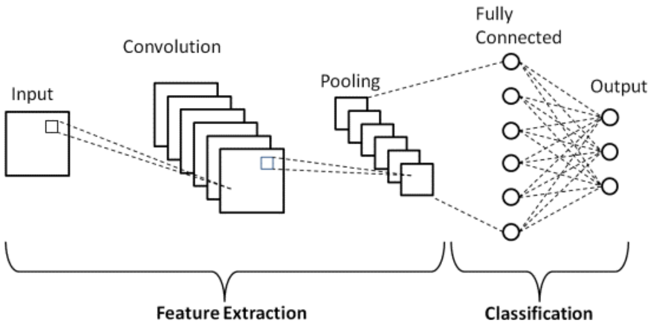


Fig 6: Convolutional Neural Network

1. The input video is first preprocessed and divided into multiple frames using frame extraction techniques.
2. Each frame is resized and normalized to match the input requirements of the CNN model.
3. The CNN model processes each frame through convolutional layers to extract visual features such as edges, textures, and facial patterns.
4. Pooling layers reduce the spatial dimensions while preserving important features.
5. The extracted features are passed through fully connected layers for classification.
6. Predictions from multiple frames are combined to analyze temporal consistency across the video.
7. The final classification determines whether the video is real or deepfake.

2.2.3. Audio Deepfake Detection using Random Forest

The audio deepfake detection module uses a Random Forest classifier, which is an ensemble learning method that combines multiple decision trees to improve classification accuracy. Deepfake audio often contains subtle inconsistencies in voice patterns, pitch variations, and frequency distributions. By extracting important acoustic features and analyzing them using multiple decision trees, the Random Forest model can effectively detect whether the audio signal is natural or artificially generated.

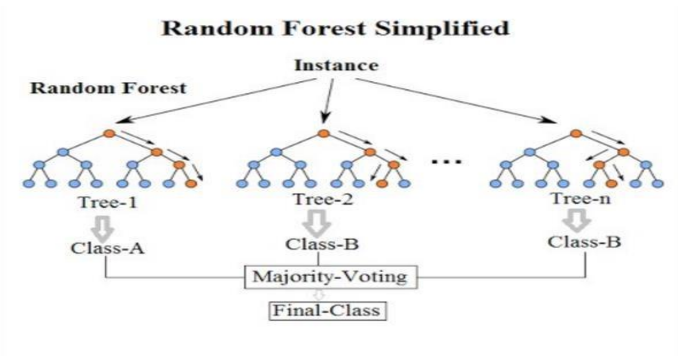


Fig 7: Random Forest

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

1. The input audio is first preprocessed to remove noise and standardize the signal format.
2. Important audio features such as MFCC, pitch, and frequency patterns are extracted from the audio signal.
3. These features are converted into numerical feature vectors.
4. The feature vectors are provided as input to the Random Forest classifier.
5. Multiple decision trees analyze different combinations of acoustic features.
6. Each decision tree produces a prediction regarding whether the audio is real or synthetic.
7. The final result is determined by majority voting among the decision trees.

2.2.4. Text Detection using RoBERTa Transformer

The textual analysis module employs the RoBERTa transformer model, which is an improved version of the BERT architecture designed for natural language understanding tasks. RoBERTa is capable of capturing deep contextual relationships between word s in a sentence using self-attention mechanisms. This allows the system to detect AI-generated or manipulated text by identifying unnatural language patterns, repetitive structures, and inconsistencies in writing style.

1. The input text is first preprocessed by removing unnecessary symbols and formatting.
2. The cleaned text is tokenized using the RoBERTa tokenizer.
3. Each token is converted into contextual embeddings. These embeddings are passed through multiple transformer layers containing self-attention mechanisms.
4. The model analyzes relationships between words and sentence structures.
5. RoBERTa detects unusual patterns that may indicate AI-generated or manipulated text.
6. The final classification layer determines whether the text is human-written or AI-generated.

2.2.5. Plagiarism Detection using Term Frequency–Inverse Document Frequency (TF-IDE)

The plagiarism detection component uses TF-IDF combined with Cosine Similarity to measure similarity between textual documents. TF-IDF calculates the importance of words within a document relative to a collection of documents. Cosine similarity then measures the similarity between document vectors. This approach helps identify duplicated or highly similar content, which may indicate plagiarism.

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

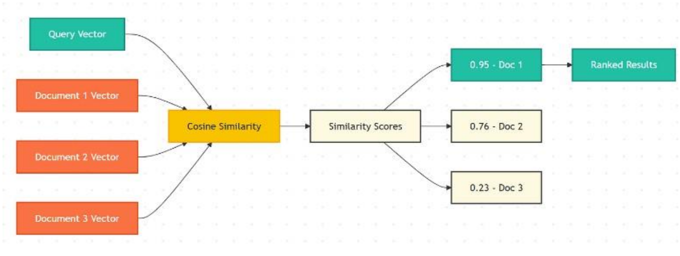


Fig 8: TF-IDE

1. The input documents are first preprocessed by removing stop words and performing tokenization.
2. Each document is converted into a TF-IDF vector representation.
3. The TF-IDF vectors represent the importance of each word in the document.
4. Cosine similarity is calculated between the vectors of different documents.
5. The similarity score indicates how closely related the documents are higher similarity score suggests possible duplication or plagiarism.
6. The system reports the similarity result for further analysis.

3. Multimodal Feature Fusion and Analysis

The suggested deepfake detection framework relies heavily on multimodal feature fusion and analysis, which integrates data from many data modalities, including text, photos, videos, and audio signals. Manipulated media may show inconsistencies across modalities in real-world situations, and depending solely on one modality may not yield enough evidence for accurate identification. In order to create a complete representation of the input media, the suggested system integrates the features that were taken from each detection model. Meaningful features pertaining to linguistic structures, speech traits, or visual patterns are extracted independently by each modality-specific model. For additional analysis, these features are combined and converted into a single feature space. The system can record complementary and correlated data from many modalities thanks to the fusion process, which increases the detection process's robustness and dependability. The technique can find tiny manipulation artifacts and cross-modal discrepancies that single-modality analysis could miss by examining the fused feature representation. The final judgment module then processes the integrated feature representation to produce a forecast that indicates whether the input media is modified or legitimate. By utilizing the advantages of several detection models, this multimodal fusion technique greatly improves the deepfake detection system's overall performance.

3. Result Generation

The system creates a final prediction that indicates whether the content is modified (fake) or legitimate (real) after proces sing the input media using the relevant detection model. The

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

matching machine learning or deep learning model for each media modali ty is used to analyze it, and the model's prediction score establishes the input's legitimacy. The SIGLIP vision transformer-based detector supplied by the API is used by the system for picture inputs. To ascertain if an image was created by artificial intelligence or obtained from real-world sources, the model examines visual patterns, textures, and semantic errors. The probability score that the API returns indicates how likely it is that the image was created by AIA Convolutional Neural Network (CNN) analyzes extracted frames from video inputs. The CNN model picks up on facial irregularities and spatial characteristics that are frequently found in deepfake films. The network determines if the video is authentic or altered based on these retrieved properties. The system uses a Random Forest Classifier trained on extracted audio features for audio inputs. To ascertain whether the audio is real or arti ficially produced, the classifier examines patterns in the voice stream, such as changes in frequency, tone, and waveform features.

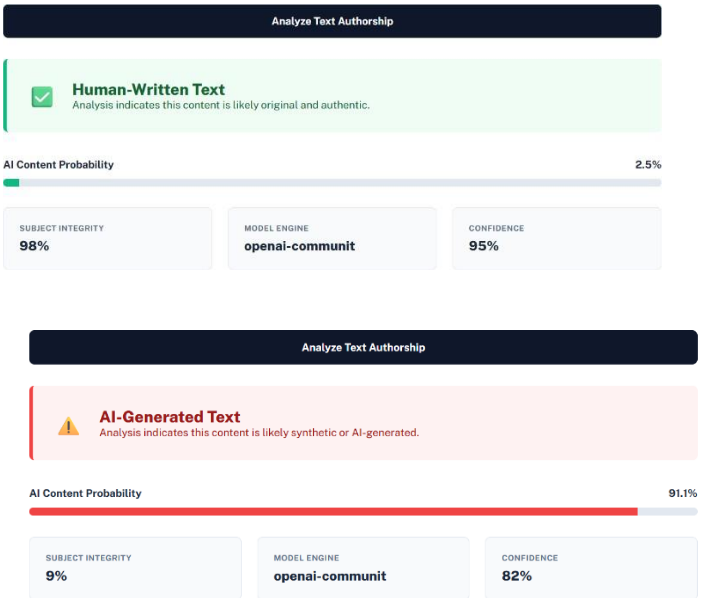


Fig 9: Audio Detection Results

The system makes use of the transformer detector API based on RoBERTa for text inputs. To determine if the content was created by artificial intelligence or authored by a person, this approach examines linguistic patterns, token distributions, and contextual structures in the text.

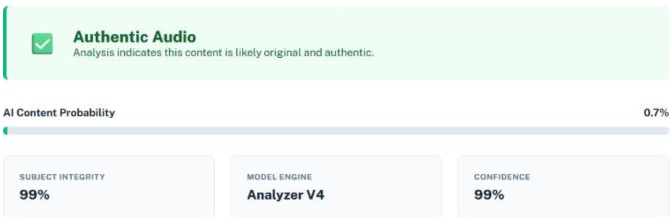
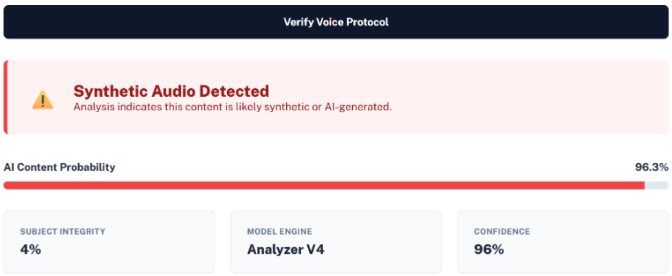


Fig 10: Text Detection Results



The technique uses cosine similarity in conjunction with TF-IDF vectorization to identify document plagiarism. Similarity scores between the input document and reference content are computed after the textual content of the document is transformed into numerical vectors. Plagiarism may be indicated by a high similarity score. The system generates the final classification result, which shows whether the input media is authentic or fraudulent, together with the related confidence score where available, once the appropriate model generates its prediction. After that, the user interface receives this outcome for interpretation and visualization.

4. Results And Discussion

Images, videos, audio files, text, and documents were among the media inputs that were tested in order to assess the suggeste d multimodal deepfake detection system. To ascertain whether the material was real or synthetic, each type of input was process ed using the appropriate machine learning or deep learning model. The system's results show that it is possible to effectively identify corrupted digital media by integrating many detection techniques.

The SIGLIP Vision Transformer-based detector API, which categorizes images as either AI-generated or human-created, was used by the system for image analysis. Visual discrepancies, false textures, and synthetic patterns that are commonly found in AI- generated photographs were correctly detected by the model. The model was able to consistently discern between generated images and real photographs, according to experimental testing. The constructed Convolutional Neural Network (CNN) examined frames taken from the input video in order to recognize video deepfakes. Spatial elements that frequently show up in edited v ideos were caught by the CNN, including facial deformities, irregular pixel distributions, and irregularities in facial regions. The outcomes demonstrated that by examining frame-level patterns, the model could successfully identify deepfake films.

The system used a Random Forest Classifier trained on extracted audio features for audio detection. The classifier assessed waveform shapes, amplitude fluctuations, and frequency patterns in speech. The findings of the experiment showed that the Random Forest model did a good job of differentiating between artificially produced or altered audio and real human speech. The system employed the RoBERTa-based transformer detection API for text detection, which examines the text's linguistic and contextual patterns. The program revealed statistical anomalies in token usage and sentence patterns that are frequently present in machine-

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

generated text. The outcomes demonstrated the model's ability to correctly identify text as either AI-generated or human-written.

The system used cosine similarity and TF-IDF vectorization to measure textual content similarity in order to detect plagiarism in documents. By calculating similarity scores across text vectors, the method proved successful in finding copied or extremely similar textual content within documents, according to the results. Overall, the experimental findings show that the suggested system may use specific models for each modality to assess various kinds of digital material. The system offers a thorough method for recognizing modified content by combining various detection approaches into a unified framework. When compared to using a single detection method, the implementation demonstrates that multimodal analysis increases the reliability of deepfake detection.

5. Future Work

Even though the suggested multimodal deepfake detection system performs well in a variety of media, there are still a number of areas that could use better. Future studies can concentrate on improving the detection framework's real-time capabilities, scalability, and accuracy.

Adding more deep learning architectures to increase detection accuracy is one such enhancement. To better capture intricate patterns found in modified media, hybrid neural network architectures and sophisticated transformer-based models should be investigated. Particularly for extremely realistic deepfakes, these models might aid in enhancing detection performance. The creation of a real-time detecting system is another crucial area for future research. Nowadays, the majority of detection models work in an offline setting where uploaded media files are examined. The technology may instantaneously detect corrupted content on social media sites, video streaming services, and communication networks by putting real-time detection algorithms into place. Expanding the dataset used to train the models could be another improvement in the future. Larger and more varied datasets can strengthen the system's resilience and enable it to identify recently developed deepfake creation methods. Maintaining detection reliability will be aided by ongoing model retraining with updated datasets.

6. Conclusion

In order to detect manipulated digital media in a variety of formats, such as photographs, videos, audio, text, and documents, a sophisticated multimodal deepfake detection system was created in this work. The suggested method uses specific detection models to assess each kind of material by integrating various machine learning and deep learning approaches. The system enhances its capacity to recognize artificial or modified content by integrating several detection algorithms into a single framework. A Vision Transformer-based model was used for picture analysis in order to differentiate between images produced by AI and those created by humans. A Convolutional Neural Network (CNN), which examines video frames to spot visual irregularities, was used to detect video deepfakes. A Random Forest classifier that analyzes acoustic characteristics to identify edited speech was used to assess audio authenticity. To ascertain whether the text was produced by artificial intelligence, a transformer model based on RoBERTa was utilized for text analysis. In order to identify possible plagiarism, document authenticity was also assessed using T F- IDF vectorization with cosine similarity. The results of the experiment show that the suggested method can successfully detect modified material in a variety of media. The technology offers a complete solution for identifying deepfakes and AI-generated content by using specific models for each modality. This strategy aids in addressing the escalating problem of media manipulation and digital disinformation.

All things considered, the created framework shows how integrating several detection methods can improve deepfake detection systems' accuracy and dependability. The suggested method can be used to guarantee the integrity and validity of digital content in fields like cybersecurity, journalism, social media monitoring, and digital forensics.

References

1. A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies and M. Nießner, "FaceForensics++: Learning to Detect Manipulated Facial Images," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019, pp. 1 -11, doi: 10.1109/ICCV.2019.00009.
2. D. Afchar, V. Nozick, J. Yamagishi and I. Echizen, "MesoNet: a Compact Facial Video Forgery Detection Network," 2018 IEEE Int ernational Workshop on Information Forensics and Security (WIFS), Hong Kong, China, 2018, pp. 1 -7, doi: 10.1109/WIFS.2018.8630761.
3. Y. Li and S. Lyu, "Exposing DeepFake Videos by Detecting Face Warping Artifacts," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Long Beach, CA, USA, 2019, pp. 46 -52, doi: 10.1109/CVPRW.2019.00020.
4. Y. Li, M. -C. Chang and S. Lyu, "In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking," 2018 IEEE International Work shop on Information Forensics and Security (WIFS), Hong Kong, China, 2018, pp. 1 -7, doi: 10.1109/WIFS.2018.8630787.
5. D. Güera and E. J. Delp, "Deepfake Video Detection Using Recurrent Neural Networks," 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Auckland, New Zealand, 2018, pp. 1 -6, doi: 10.1109/AVSS.2018.8639163.
6. C. N. Duong, K. G. Quach, I. Jalata, N. Le and K. Luu, "MobiFace: A Lightweight Deep Learning Face Recognition on Mobile Devi ces," 2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS), Tampa, FL, USA, 2019, pp. 1 -6, doi: 10.1109/BTAS46853.2019.9185981.
7. J. Thies, M. Zollhöfer, M. Stamminger, C. Theobalt and M. Nießner, "Face2Face: Real-Time Face Capture and Reenactment of RGB Videos," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 2387 -2395, doi: 10.1109/CVPR.2016.262.
8. Y. Nirkin, Y. Keller and T. Hassner, "FSGAN: Subject Agnostic Face Swapping and

Reenactment," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019, pp. 7183 -7192, doi: 10.1109/ICCV.2019.00728.

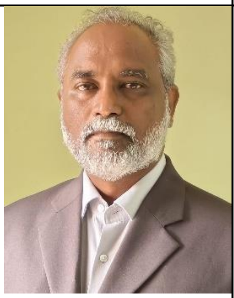
9. K. P. Nandan, B. Pakruddin, S. Afridi, S. K. R. and S. V. Pati, "Real-Time Detection of Fake News Articles Using Deep Learning Techniques," 2025 International Conference on Next Generation Communication & Information Processing (INCIP), Bangalore, India, 2 025, pp. 687-691, doi: 10.1109/INCIP64058.2025.11019208.

10. V. C. Handaragall and J. Subhash Metan, "Comprehensive Analysis of Deep Fake Detection Techniques for Images and Videos using Deep Learning," 2025 International Conference on Intelligent Computing and Knowledge Extraction (ICICKE), Bengaluru, India, 2025, pp. 1-5, doi: 10.1109/ICICKE65317.2025.11136286.

Biographies of Authors

	**Kumpatla Chinna** currently residing at main road, challapalli (533213), is a B. Tech student specializing in Artificial Intelligence and Machine Learning at Bonam Venkata Chalamayya Engineering College, Odalarevu, with an expected graduation in April 2026. He aims to secure a position that leverages his strong organizational skills, educational background, and ability to work effectively with others. He possesses key skills in communication, self-motivation, analytical and critical thinking, and dependability. While his professional experience is listed as a student, his proactive approach and skill set indicate a strong potential for growth and contribution in a professional setting. For further contact, he can be reached at 8184936555 or via email at kumpatlachinna22@gmail.com
	**Bolla Siva Naga kumar** currently residing at 4-114, tarakarama Veedi, Manepalli (533249), is a B. Tech student specializing in Artificial Intelligence and Machine Learning at Bonam Venkata Chalamayya Engineering College, Odalarevu, with an expected graduation in April 2026. He aims to secure a position that leverages his strong organizational skills, educational background, and ability to work effectively with others. B Siva Naga kumar possesses key skills in communication, self-motivation, analytical and critical thinking, and dependability. While his professional experience is listed as a student, his proactive approach and skill set indicate a strong potential for growth and contribution in a professional setting. For further contact, he can be reached at 7993544599 or via email at bsivanagakumar09@gmail.com.
	**Somisetti Satya** is a B. Tech student specializing in Artificial Intelligence and Machine Learning at Bonam Venkata Chalamayya Engineering College, Odalarevu, with an expected graduation in April 2026. She aims to secure a position that leverages his strong organizational skills, educational background, and ability to work effectively with others. Satya possesses key skills in communication, self-motivation, analytical and critical thinking, and dependability. While his professional experience is listed as a student, his proactive approach and skill set indicate a strong potential for growth and contribution in a professional setting. For further contact, he can be reached at 8919260158 or via email at satyasomisetti111@gmail.com

International Journal of Science and Engineering Science Research
Volume 2 | Issue 1 | January-March 2026 | ISSN 3049-2793
Peer Reviewed (Referred Journal)
DOI: https://doi.org/10.63671/ijsesr.v2i1.70

	**Polisetti sri surya mahesh** currently residing at challapalli,, is a B. Tech student specializing in Artificial Intelligence and Machine Learning at Bonam Venkata Chalamayya Engineering College, Odalarevu, with an expected graduation in April 2026. He aims to secure a position that leverages his strong organizational skills, educational background, and ability to work effectively with others. He possesses key skills in communication, self-motivation, analytical and critical thinking, and dependability. While his professional experience is listed as a student, his proactive approach and skill set indicate a strong potential for growth and contribution in a professional setting. For further contact, he can be reached at 83417 01608 or via email at polisettisurya16@gmail.com
	**Bokka Satish** a B. Tech student specializing in Artificial Intelligence and Machine Learning at Bonam Venkata Chalamayya Engineering College, Odalarevu, with an expected graduation in April 2026. He aims to secure a position that leverages his strong organizational skills, educational background, and ability to work effectively with others. He possesses key skills in communication, self-motivation, analytical and critical thinking, and dependability. While his professional experience is listed as a student, his proactive approach and skill set indicate a strong potential for growth and contribution in a professional setting. For further contact, he can be reached at 80967 80014 or via email at satishbokka102@gmail.com
	**Alapati Ramakrishna** is a Professor in the Department of Artificial Intelligence and Machine Learning at Bonam Venkata Chalamayya Engineering College, located in Odalarevu, Andhra Pradesh, India. He resides in Gandhinagar, Amalapuram. He is actively involvedinteachingand researchinthefield of Artificial Intelligenceand MachineLearning. Throughhis academic experience, he contributes to the development of innovative technologies and student learning. He is associated with BVC Group of Institutions, supporting academic growth and research activities. He guides students in projects, research work, and technical development. His dedication helps in promoting advanced computing knowledge among engineering students. He can be contacted at 9866946554 or through email at ark.bvce@bvcgroup.in