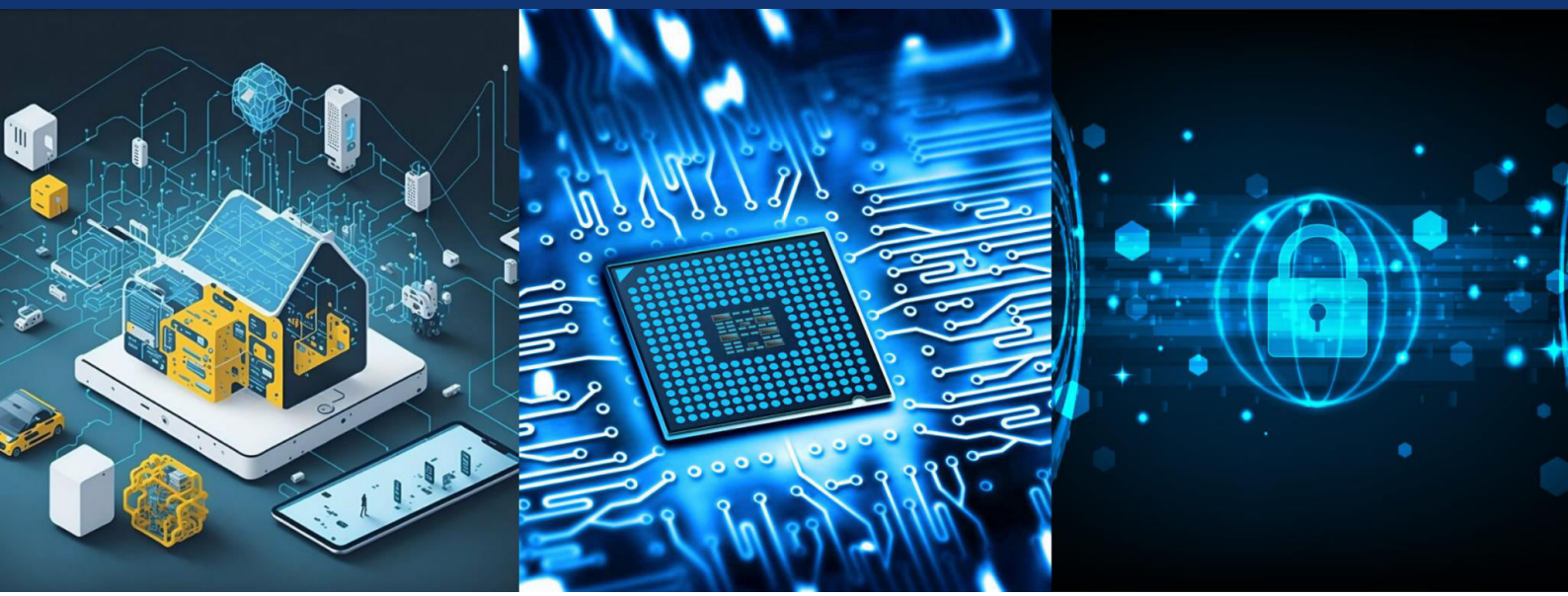
IJIRCCE

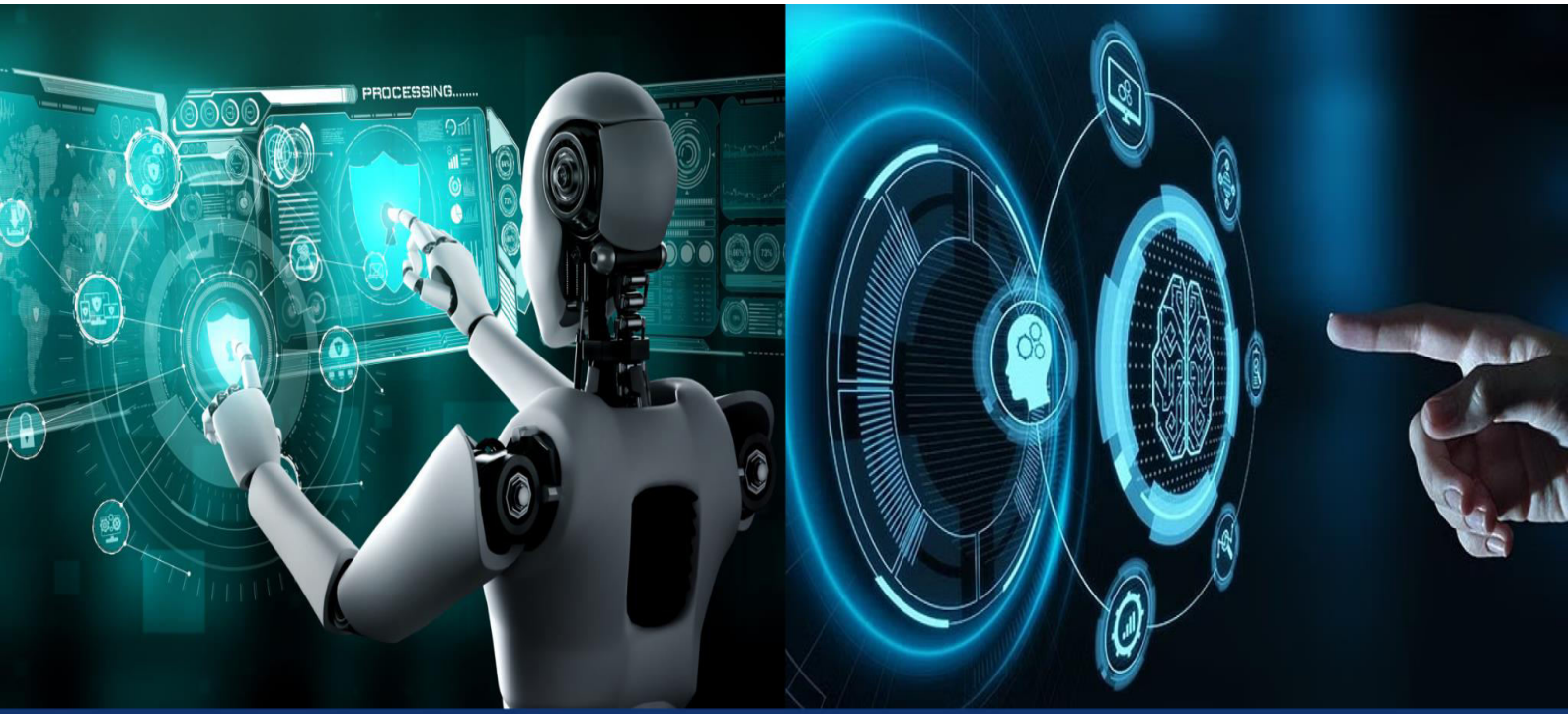
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.927

Volume 14, Issue 5, May 2026

© 2026 IJIRCCE | Volume 14, Issue 5, May 2026 DOI: 10.15680/IJIRCCE.2026.1405091

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Digital Doppelgangers

Manoj Kumar[1], Lakshmi Narasimha[2], Dr. K. Srikala[3], Ms. Dr. K Rajitha[4], Ms. K. Shirisha[5], Mr. Manas Rath[6]

Undergraduate Students, Department of Computer Science and Engineering, Mahatma Gandhi Institute of Technology, Hyderabad, Telangana, India[1-2]

Assistant Professors, Department of Computer Science and Engineering, Mahatma Gandhi Institute of Technology, Hyderabad, Telangana, India[3-6]

ABSTRACT: Based on the present globalization, detection of deepfakes is one of the largest challenges in preserving the integrity of digital media. This project designs a deepfake detection system, where a convolutional neural network (CNN) is implemented using TensorFlow and supervised learning to classify images as either "Real" or "Fake." The system has a Deepfake Detector Model and a Dataset Handler, as well as a Flask-based web application that does real-time predictions. The architecture used in the model includes convolutional and pooling layers for extracting features and then dense layers for classification. The train will include preprocessing of the data rescaling and normalization together with callbacks, such as early stopping and learning rate reduction, for better performance. Validating effectiveness, its measures consist of accuracy, precision, and recall. To generate a real fake image, generative adversarial networks integrate well within, augmenting the size of the dataset to improve training robustness. Dataset Handler deals with automating download, unzipping, and data processing. This saves the trained model and puts it in the Flask application, where users can upload images and get their respective classifications. The above implementation explains why TensorFlow-based CNNs have such significant efficacy against GAN-generated deepfake content.

KEYWORDS: Deepfake Detection, Convolutional Neural Network (CNN), Machine Learning, Image Processing, TensorFlow, Flask, Computer Vision, Feature Extraction, Binary Classification, Model Training, Data Preprocessing, GAN, Artificial Intelligence, Image Classification, Real-Time Predictio

I. INTRODUCTION

Deepfakes—AI-generated images, audio, and videos with highly realistic manipulation—have become a serious concern in today's digital world. Synthetic media poses threats to security, privacy, and trust across platforms such as social media and news outlets by enabling misinformation, identity theft, and fraud. As deepfake generation tools, particularly those using GANs, grow more advanced, distinguishing real content from fabricated media has become increasingly difficult. Therefore, there is a critical need for effective detection mechanisms. Deepfake detection is a challenging task that requires identifying subtle differences between real and synthetic media. Traditional methods rely on detecting visual artifacts or inconsistencies such as abnormal facial expressions, lighting, or textures, but their effectiveness decreases as deepfake quality improves. Convolutional Neural Networks (CNNs) have proven to be highly effective in detecting deepfakes by learning patterns from large datasets. This project leverages TensorFlow to build a CNN-based model that classifies images as real or fake. Additionally, Generative Adversarial Networks (GANs) are used to generate synthetic images, enriching the dataset and improving model robustness. Incorporating GAN-generated data enhances the system's ability to detect diverse deepfake types. The project also aims to develop a scalable, real-time detection system integrated with a web application for user interaction. By combining deep learning and adversarial training, this approach addresses real-world challenges posed by deepfake technologies. As deepfakes continue to evolve, developing advanced detection tools becomes increasingly important. This work contributes toward building a safer and more trustworthy digital environment.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. LITERATURE SURVEY

[1] V. Patel et al., "Deepfake Detection Using Deep Learning: A Survey," International Journal ofScientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT),vol. 12, no. 1, 2026.
[2] B. G. Deepa et al., "An enhanced deep learning framework for DeepFake detectionusingEfficientNet-B3 comparative evaluation of deep and machine learning techniques," DiscoverComputing, 2026.
[3] A. J. Xu and B. Chen, "Classifying DeepfakesUsingSwinTransformers," arXiv preprint arXiv:2501.XXXXX, 2025.
[4] Hussain,"RobustMultimodal DeepfakeDetectionviaSAFF+CMGAN," Frontiers in BigData, 2025.
[5] F. T. Winata et al., "Comparison of Deepfake Detection Using CNN and Hybrid Models," ProcediaComputer Science, Elsevier, 2025.
[6] Dharma et al., "Deep-fake Image DetectionUsingEnsemble Method," G-Tech: Jurnal, 2025
[7] M. S. Sheikh et al., "AI-Powered Deepfake Detection Using CNN and Vision TransformerArchitectures,"in Proceedings of IEEE Conference, 2024.
[8] Lanzino et al., "Faster Than Lies: Real-time Deepfake Detection Using Binary Neural Networks," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),2024.
[9] Pellicer et al., "PUDD: Multimodal Prototype-based Deepfake Detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024.
[10] Tan et al., "Rethinking Up-Sampling in CNN-Based Deepfake Detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024.

2.1. DESIGN METHODOLOGY OF PROPOSED SYSTEM

1. Data Collection Layer

The system starts by collecting datasets containing both real and fake images. Fake images are generated using GANs, while real images are obtained from authentic sources. This improves diversity and model generalization.

2. Data Preprocessing Layer

Collected images are cleaned, resized, normalized, and augmented. Techniques like cropping, flipping, and noise addition are applied to enhance robustness and handle variations.

3. Feature Extraction Layer

Preprocessed images are passed into a Convolutional Neural Network (CNN), which extracts key features such as facial structures, textures, and inconsistencies.

4. Model Training Layer

The CNN model is trained using the processed dataset on TensorFlow. The system learns patterns through multiple epochs and improves performance using backpropagation.

5. Classification Layer

The trained model classifies images into **Real** or **Fake** categories and generates a probability score indicating prediction confidence.

6. Adversarial Training Layer

GAN-generated fake images are continuously fed back into the training process to improve detection capability against new and evolving deepfakes.

7. Deployment Layer

The trained model is deployed as a backend service capable of processing user inputs in real time, ensuring scalability and efficiency.

8. User Interface Layer

A web-based interface allows users to upload images and instantly receive detection results in a simple and user-friendly manner.

9. Output & Monitoring Layer

The system displays classification results along with confidence scores and maintains logs for performance tracking, analysis, and future improvements.

© 2026 IJIRCCE | Volume 14, Issue 5, May 2026 DOI: 10.15680/IJIRCCE.2026.1405091

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

2.2. SYSTEM FLOW DIAGRAM

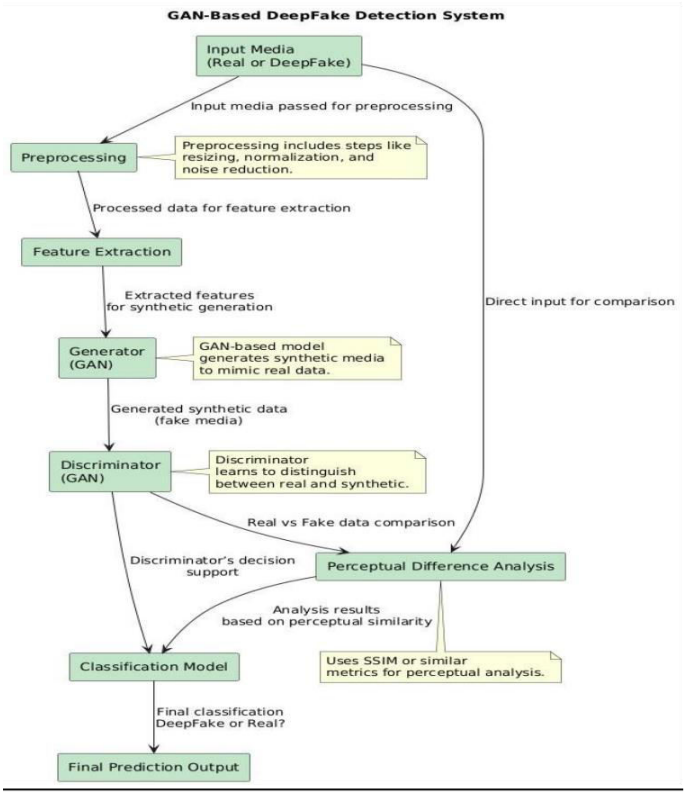


Figure 3.2 Process flow diagram

Figure 3.2 The system can be explained using a **flow diagram representation**, where the overall function is expressed as f(media)= Final Prediction. The process starts with the input media, which may be either real or deepfake, and this data is passed into the system for further processing. Initially, the media undergoes preprocessing, where important steps such as resizing, normalization, and noise reduction are applied to improve data quality and ensure consistency. This stage prepares the input so that meaningful patterns can be extracted effectively in the next step.

Once preprocessing is completed, the refined data flows into the feature extraction stage, where key features and patterns are identified from the media. These extracted features are then forwarded to the GAN generator, which creates synthetic (fake) data that mimics real-world media. The generated data, along with the original processed input, is then passed to the discriminator. The discriminator plays a critical role by learning to distinguish between real and generated (fake) data through continuous comparison, thereby improving the model's detection capability over time.

In parallel, the system also performs perceptual difference analysis, where both the original input and generated data are compared using similarity metrics such as structural similarity (SSIM). This analysis helps in identifying subtle visual differences that may not be easily detectable, providing additional support to the detection process. The outputs from the discriminator and perceptual analysis are then combined and sent to the classification model, which makes the final decision based on all the processed information.

Finally, the classification model produces the output in the form of a prediction, indicating whether the given media is real or a deepfake. This completes the flow of data through the system, moving step by step from input acquisition, preprocessing, feature extraction, GAN-based generation and discrimination, perceptual analysis, and final classification, ensuring an accurate and reliable deepfake detection process.

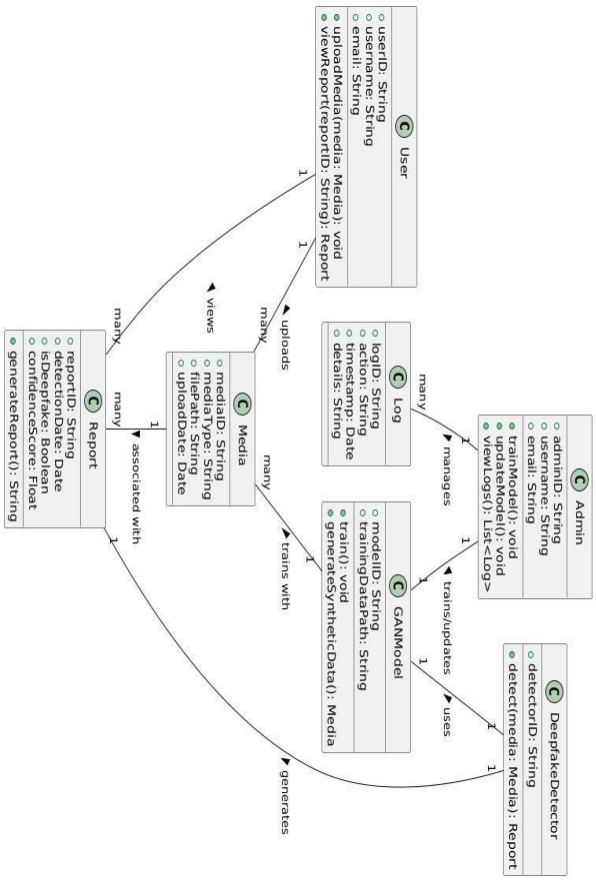
© 2026 IJIRCCE | Volume 14, Issue 5, May 2026 DOI: 10.15680/IJIRCCE.2026.1405091

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

III. BLOCK DIAGRAM



3.1 TESTING AND RESULTS

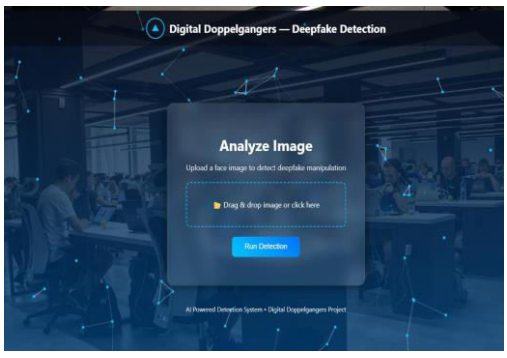


Figure 3.1.1: Deepfake Detection Interface (Input Screen)

Figure 4.1.1 This screen represents the main user interface of the system titled **"Digital Doppelgangers — Deepfake Detection."** It is designed to allow users to easily upload an image for analysis. The interface contains a central panel labeled **"Analyze Image"**, where users can either drag and drop an image or click to browse files from their system. The clean and modern UI improves user experience and makes the system accessible even for non-technical users. The background design with network-like patterns visually represents artificial intelligence and deep learning concepts. Once the user uploads an image, they can click the **"Run Detection"** button to initiate the deepfakedetection process. This screen acts as the entry point where raw input data is provided to the system

© 2026 IJIRCCE | Volume 14, Issue 5, May 2026 DOI: 10.15680/IJIRCCE.2026.1405091

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

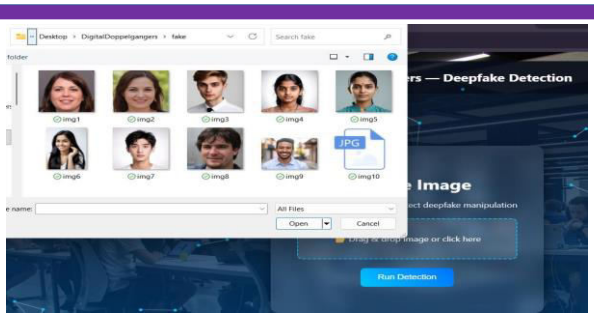


Figure 3.1.2: Image Selection (Dataset/Input Upload)

This screen shows the file selection window where the user chooses an image from their local system. The dataset folder contains multiple face images (e.g., img1, img2, etc.), which are used either for testing or validation. This step is important because it demonstrates how the system accepts real-world data as input. The images are typically preprocessed later (resizing, normalization) before being passed to the trained model. This stage bridges the gap between user interaction and backend processing, ensuring that the correct input is selected for analysis.

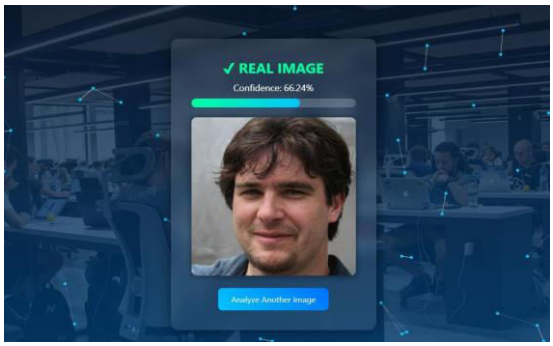


Figure 3.1.3: Prediction Result – Real Image

This screen displays the output when the system detects that the uploaded image is real (authentic). The result is clearly shown with a green indicator and a label "REAL IMAGE", making it easy to understand. Along with the prediction, the system provides a confidence score (54.83%), which indicates how certain the model is about its decision. A progress bar visually represents this confidence level. The analyzed image is also displayed below the result for verification. Additionally, there is an option "Analyze Another Image", allowing users to quickly test more images. This output confirms that the model has successfully identified genuine content.

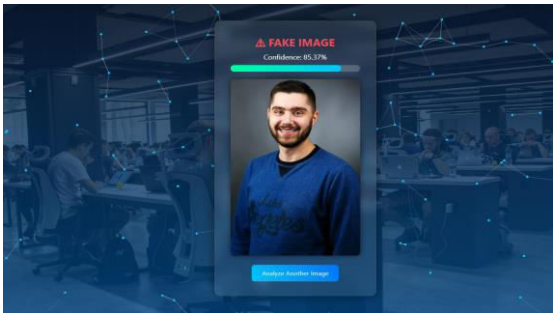


Figure 3.1.4: Prediction Result – Fake Image

This screen shows the output when the system detects a **deepfake (manipulated image)**. The result is highlighted with a red warning indicator and the label **"FAKE IMAGE"**, clearly distinguishing it from real results. The confidence score

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

(85.37%) is higher in this case, indicating strong certainty in detection. Similar to the previous screen, a progress bar visually represents the confidence level, and the analyzed image is displayed for reference. This output demonstrates the system's ability to accurately identify manipulated or synthetic media. The presence of a retry option (**"Analyze Another Image"**) ensures continuous usability.

REFERENCES

[1] T. Wang, X. Liao, K. P. Chow, X. Lin, and Y. Wang, "Deepfake Detection: A Comprehensive Survey from the Reliability Perspective," arXiv preprint arXiv:2211.10881, 2022.
Available: https://arxiv.org/abs/2211.10881
[2] G. Pei, H. Li, Y. Zhang, X. Chen, and J. Liu,"Deepfake Generation and Detection: A Benchmark and Survey," arXiv preprint arXiv:2403.17881, 2024.
Available: https://arxiv.org/abs/2403.17881
[3] L. Lin, X. He, Y. Ju, X. Wang, and F. Ding,
"Preserving Fairness Generalization in Deepfake Detection," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2024.
Available:
https://openaccess.thecvf.com/content/CVPR2024/html/Lin_Preserving_Fairness_Generalization_in_Deepfake_Detection_CVPR_2024_paper.html
[4] Z. Yan, H. Zhang, Y. Liu, X. Chen, and J. Wang, "Transcending Forgery Specificity with Latent Space Augmentation for Generalizable Deepfake Detection," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2024.
Available:
https://openaccess.thecvf.com/content/CVPR2024/html/Yan_Transcending_Forgery_Specificity_with_Latent_Space_Augmentation_for_Generalizable_Deepfake_CVPR_2024_paper.html
[5] J. Choi, M. Kim, S. Lee, H. Park, and Y. Jung,"Exploiting Style Latent Flows for Generalizing Deepfake Video Detection," arXiv preprintarXiv:2403.06592, 2024.
Available: https://arxiv.org/abs/2403.06592
[6] M. Li, Q. Zhao, H. Sun, X. Wang, and J. Li,"A Survey on Speech Deepfake Detection," arXiv preprint arXiv:2404.13914, 2024.
Available: https://arxiv.org/abs/2404.13914
[7] F. A. Croitoru, A. Serban, M. Popescu, I. Ionescu, and D. Marinescu,"Deepfake Media Generation and Detection in the Age of Generative AI," arXiv preprint arXiv:2411.19537, 2024.
Available: https://arxiv.org/abs/2411.19537
[8] H. Nguyen-Le, T. Nguyen, P. Tran, Q. Le, and V. Pham,"Passive Deepfake Detection Across Multi-modalities," arXiv preprint arXiv:2411.17911, 2024.
Available: https://arxiv.org/abs/2411.17911
[9] P. Liu, Q. Tao, J. T. Zhou, X. Zhang, and L. Wu,"Evolving from Single-modal to Multi-modal Facial Deepfake Detection: A Survey," arXiv preprint arXiv:2406.06965, 2024.
Available: https://arxiv.org/abs/2406.06965
[10] M. Alrashoud, A. Khan, S. Malik, R. Ahmed, and T. Hussain, "Deepfake Video Detection Methods, Approaches, and Trends," Expert Systems with Applications, 2025.
Available: https://www.sciencedirect.com/science/article/pii/S111001682500465X

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details